

\*\*\*\*\*

WQEQELH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:18:15 1998; MasPar time 17.49 Seconds  
Tabular output not generated. 739.991 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHNRIISVLLTLLAVV.....ADGNGFTVNGGAVSVVVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot35  
1:swiss1

Statistics: Mean 50.716; Variance 89.551; scale 0.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3368	87.0	518	1	AMT6_BACS7	0.00e+00
2	2679	69.2	512	1	AMY_BACLI	0.00e+00
3	2645	68.3	549	1	AMY_BACST	0.00e+00
4	2559	66.1	514	1	AMY_BACAM	0.00e+00
5	1403	36.2	494	1	AMY2_SALTY	0.00e+00
6	1360	35.1	495	1	AMY2_ECOLI	1.24e-289
7	217	5.6	548	1	AMT4_PSESA	1.98e-279
8	218	5.6	551	1	AMT4_PSESA	2.65e-20
9	202	5.2	528	1	AMY_BACCI	1.70e-20
10	192	5.0	713	1	AMYR_BACS8	1.81e-17
11	189	4.9	712	1	CDGT_BACS3	1.30e-15
12	189	4.9	718	1	CDGT_BACSS	4.62e-15
13	185	4.8	713	1	CDGU_BACCI	4.62e-15
14	186	4.8	718	1	CDGT_BACCI	2.48e-14
15	182	4.7	711	1	CDGT_BACST	1.63e-14
16	183	4.7	713	1	CDGT_BACS0	8.67e-14
17	183	4.7	713	1	CDGT_BACSP	5.71e-14
18	181	4.7	718	1	CDGT_BACLI	5.71e-14
19	172	4.4	413	1	AMY3_WHEAT	1.31e-13
20	158	4.1	421	1	AMYA_VITGM	5.32e-12
21	147	3.8	713	1	CDG2_BACMA	1.45e-09
22	147	3.8	919	1	AMY_STRLI	1.03e-07
23	145	3.7	714	1	CDGI_BACMA	1.03e-07
					CYCLONALTODEXTRIN	2.20e-07

24	138	3.6	428	1	AMY1_ORISA	ALPHA-AMYLASE PRECURSO	3.02e-06
25	140	3.6	440	1	AM3A_ORISA	PROBABLE ALPHA-AMYLASE	1.44e-06
26	140	3.6	581	1	AMY1_SCHPO	ALPHA-AMYLASE PRECURSO	2.08e-06
27	139	3.6	710	1	AMY_THETU	BETA-AMYLASE (EC 3.2.1.1)	3.02e-06
28	138	3.6	1196	1	AMYB_BACPO	ALPHA-AMYLASE TYPE B I	9.08e-06
29	135	3.5	135	1	AMY5_HORVU	ALPHA-AMYLASE TYPE B I	4.36e-06
30	137	3.5	368	1	AMY3_HORVU	ALPHA-AMYLASE TYPE B I	4.36e-06
31	137	3.5	427	1	AMY2_HORVU	ALPHA-AMYLASE TYPE B I	4.36e-06
32	135	3.5	435	1	AM3D_ORISA	ALPHA-AMYLASE ISOZYME	6.30e-06
33	136	3.5	443	1	AM2C_ORISA	ALPHA-AMYLASE ISOZYME	6.30e-06
34	136	3.5	445	1	AMC2_ORISA	ALPHA-AMYLASE ISOZYME	6.30e-06
35	133	3.4	429	1	AMY6_HORVU	ALPHA-AMYLASE TYPE A I	1.88e-05
36	132	3.4	438	1	AMY1_HORVU	ALPHA-AMYLASE TYPE A I	2.70e-05
37	126	3.3	153	1	AMY4_HORVU	ALPHA-AMYLASE TYPE B I	2.28e-04
38	127	3.3	498	1	AMYA_ASAPW	ALPHA-AMYLASE A PRECUR	1.60e-04
39	127	3.3	499	1	AMYB_ASAPW	ALPHA-AMYLASE B PRECUR	1.60e-04
40	127	3.3	499	1	AMYB_ASAPW	ALPHA-AMYLASE B PRECUR	1.60e-04
41	129	3.3	499	1	AMY_ASPSH	ALPHA-AMYLASE PRECURSO	1.60e-04
42	127	3.3	717	1	AMYM_BACST	MALTOGENIC ALPHA-AMYL	7.89e-05
43	128	3.3	1592	1	GTFL_STRDO	GLUCOSYLTRANSFERASE-1	1.13e-04
44	128	3.3	1597	1	GTFL_STRDO	GLUCOSYLTRANSFERASE-1	1.13e-04
45	124	3.2	437	1	AM3E_ORISA	ALPHA-AMYLASE ISOZYME	4.58e-04

ALIGNMENTS

RESULT	1	STANDARD;	PRT;	518 AA.
ID	AMT6_BACS7			
AC	P19571			
DT	01-FEB-1991 (REL. 17, CREATED)			
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)			
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)			
DE	GLUCAN 1,4-ALPHA-MALTOHEXAOSIDASE PRECURSOR (EC 3.2.1.98) (G6-AMYLASE)			
DE	(MALTOHEXAOSIDE-PRODUCING AMYLASE) (EXO-MALTOHEXAOSIDOLASE).			
OS	BACILLUS SP. (STRAIN 707).			
OC	PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 34-36.			
RX	MEDLINE; 88162814.			
RA	TSUKAMOTO A., KIMURA K., ISHII Y., TAKANO T., YAMANE K.;			
RL	BIOCHEM. BIOPHYS. RES. COMMUN. 151:25-31(1988).			
CC	- - CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC LINKAGES			
CC	IN AMYLACEOUS POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE			
CC	MALTOHEXAOSIDE RESIDUES FROM THE NON-REDUCING CHAIN ENDS.			
CC	- - PATHWAY: DEGRADATION OF STARCH.			
CC	- - SUBCELLULAR LOCATION: EXTRACELLULAR.			
CC	- - SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO			
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.			
DR	EMBL; M18862; G142497; -			
DR	PIR; A27705; A27705.			
KW	HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.			
FT	SIGNAL 1 33			
FT	CHAIN 34 518			
FT	SEQUENCE 518 AA; 59009 MW; FB2B65AE CRC32;			

Query Match 87.0%; Score 3368; DB 1; Length 518;  
Best Local Similarity 82.8%; Pred. No. 0.00e+00;  
Matches 427; Conservative 57; Mismatches 30; Indels 2; Gaps 2;

Db	3	MRTGKGFSLTLLAFLLVITSI-PTFLVDVEAHNGTNGTMMQYFEWYLPDNGNHNRLN	61
Qy	1	MKLHNR-IIISVLLTLLAVAVLFPYTPAQAHNGTNGTMMQYFEWHLPDNGNHNRLR	59
Db	62	SDAENLSKGTITAVWIPPAWKASONDYGYGAYDLYDLGEFNKGKVTYKGYTRSOLOQA	121
Qy	60	DDAENLSKGTITAVWIPPAWKGTISQNDYGYGAYDLYDLGEFNKGKVTYKGYTRSOLOQA	119
Db	122	VTSLKNGIQYGVVNMHKGADATEMVAVEVNPNNRQNEVTGYTTEATRFDPGGR	181
Qy	120	VTSLKNGIQYGVVNMHKGADGTEMVAVENVRNQRNEISGYTTEATWKDFPGR	179
Db	182	GNTHSSFKRWYHFDGVDWDOSRLNRIYFRHGKAWDEVDTEGNGWYQYLMYADTDM	241

QY 180 GNTHSNFKRWTHFDGTDQSRQLQNKIYKFRGTGKAWDEVDIENGNDYDLYMYADIDM 239  
 Db 242 DHPVVELNLRNGVYNTNLGLDGRFRIDAVKHKISYTRDWINHRSATCKNMFVAEAFW 301  
 QY 240 DHPVVELNLRNGVYNTNLGLDGRFRIDAVKHKISYTRDWINHRSATCKNMFVAEAFW 299  
 Db 302 KNDLGAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVORHPSHAYTF 361  
 QY 300 KNDLGAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVORHPSHAYTF 359  
 Db 362 VDNHDSQPEALESEFVEEMFKPLAYALTLTREGQPSVYFYGYGIPTHGVPAMRSKIDP 421  
 QY 360 VDNHDSQPEALESEFVQSFVKPLAYALTLTREGQPSVYFYGYGIPTHGVPAMRSKIDP 419  
 Db 422 ILKARQYAYGKNDYLDHNNHIGTREGTAHPNSGLATIMSDGAGSKWFMVGRNK 481  
 QY 420 LQARQYAYGQHDFYFDHDIIGTREGDSHPNSGLATIMSDGAGSKWFMVGRNK 479  
 Db 482 QVMSDITNRTGTVINADGNFSGVSGSIWVN 517  
 QY 480 QVMSDITNRTGTVINADGNFSGVSGSIWVN 515

RESULT 2  
 ID AMY\_BACLI STANDARD; PRT; 512 AA.  
 AC P06278;  
 DT 01-JAN-1988 (REL. 06, CREATED)  
 DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYS OR AMYL.  
 OS BACILLUS LICHENIFORMIS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN-ATCC 27811;  
 RX MEDLINE; 86111694.  
 RA YUKI T., NOMURA T., TEZUKA H., TSUBOI A., YAMAGATA H.,  
 RA TSUKAGOSHI N., UDARA S.;  
 RL J. BIOCHEM. 98:1147-1156(1985).  
 RN [2] SEQUENCE FROM N.A.  
 RP MEDLINE; 86195857.  
 RX GRAY G.L., MAINZER S.E., REY M.W., LAMSA M.H., KINDLE K.L.,  
 RA CARMONA C., REQUAD C.;  
 RL J. BACTERIOL. 166:635-643(1986).  
 RN [3] SEQUENCE OF 1-104 FROM N.A.  
 RP MEDLINE; 84185455.  
 RX STEPHENS M.A., ORTLEPP S.A., OLLINGTON J.F., MCCONNELL D.J.;  
 RL J. BACTERIOL. 158:369-372(1984).  
 RN [4] SEQUENCE OF 1-29 FROM N.A.  
 RP MEDLINE; 89213924.  
 RX LAOIE B.M., CHAMBLISS G.H., MCCONNELL D.J.;  
 RL J. BACTERIOL. 171:2435-2442(1989).  
 RN [5] SEQUENCE OF 30-47.  
 RP MEDLINE; 82098050.  
 RX KUHN H., FIETZKE P.P., LAMPEN J.O.;  
 RL J. BACTERIOL. 149:372-373(1982).  
 RN [6] X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RP STRAIN-ATCC 27811;  
 RX MEDLINE; 95182462.  
 RA MACHUS M., WIEGAND G., HUBER R.;  
 RL J. MOL. BIOL. 246:545-559(1995).  
 CC -1- CATALYTIC ACTIVITY: ENDOLYSIN OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; X03236; G39552; -.

DR EMBL; M38570; G142480; -  
 DR EMBL; M13256; G142511; -  
 DR EMBL; K01984; G142433; -  
 DR EMBL; M26412; G516590; -  
 DR EMBL; A17930; G512328; -  
 DR PIR; A00844; ALBSL.  
 DR PIR; B24549; B24549.  
 DR PIR; A26151; A26151.  
 DR PDB; 1BPL; 17-AUG-96.  
 DR PDB; 1VJS; 12-MAR-97.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL; 3D-STRUCTURE.  
 FT CHAIN 1 29  
 FT SIGNAL 30 512  
 FT ACT\_SITE 260 260  
 FT ACT\_SITE 264 284  
 FT ACT\_SITE 357 357  
 FT CONFLICT 38 38  
 FT CONFLICT 163 163  
 FT CONFLICT 339 339  
 FT CONFLICT 349 349  
 SQ SEQUENCE 512 AA; 58549 MW; 4B5D334D CRC32;  
 Query Match 69.2%; Score 2679; DB 1; Length 512;  
 Best Local Similarity 66.9%; Pred. No. 0.00e+00;  
 Matches 347; Conservative 86; Mismatches 76; Indels 10; Gaps 7;  
 Db 1 MKQOKRLYARLLTLLFALIFLPH-SAAAAA--N-LNGTLMQYFEMYPNDGQHKRLQN 56  
 QY 1 MKNLRIISVLLTLLAVAVLPFYPMTPEQAHHNGTNGTMMQYFEWHLPNDGNHNRURD 60  
 Db 57 DSAVLAEGITAVWIPAYKGTSDQVGYGAYDYLGLGEFHQKGTVRKYTKGELQSAI 116  
 QY 61 DAANLASKGITAWIPAKGTSQNDVGYGAYDYLGLGEFNQKGTVRKYTKGELQSAI 120  
 Db 117 KSLHSRDINVGVDVINHKGGADATVAVEVDPADNRVVISGEHRKAWTHFFPGRG 176  
 QY 121 TSLKNGIOVYGDVYVNNHKGADGTEMVAVENVRNNOEISEGYTIEAWTKFDFPGRG 180  
 Db 177 STYDFKWHYHFDGTDWDSRKL-NRIYFQG--KAWDEVSNEGNNDYDLYMYADIDVD 233  
 QY 181 NTHSNFKRWTHFDGTDQSRQLQNKIYKFRGTGKAWDEVDIENGNDYDLYMYADIDVD 240  
 Db 234 HPDVAEIKRWGTWYANELQIDGFRDRAVAKHIFSELRDWNHVRNREKTKEMFTVAEYQ 293  
 QY 241 HPEVINELRWGWYNTNLNLDGFRIDAVKHKISYTRDWINHRSATCKNMFVAEAFW 300  
 Db 294 NDLGALENYLNKTNHNSVDFVPLHYNQFHAASSTQGGYDMRKLNLNYSVSKHPLKAVTFV 353  
 QY 301 NDLAALNENYLNKTNHNSVDFVPLHYNLNYSNNGSGYFDMRNLNLSGVVQKHPHATFV 360  
 Db 354 DNHDTPQGSLESTVQTFKPLAYAFILTRSGYPOVYFYGYDMYGTGKDSORETPALKHKI 413  
 QY 361 DNHDSPGEALESEFVQSFVKPLAYALTLTREGQPSVYFYGYGI--PTH-GVPSMKSKI 417  
 Db 414 EPILKARKOYAGQHDYFDHDIIGTREGDSVNSGLAALITDGPAGKMYVGRN 473  
 QY 418 DPLQARQYAYGQHDFYFDHDIIGTREGDSHPNSGLATIMSDGPGGKMYVGRN 477  
 Db 474 AGETHWDTGNRSEPVYINSEGEHFNHNSVDFVPLHYNLNYSNNGSGYFDMRNLNLSGVVQKHPHATFV 512  
 QY 478 AGQWRDITNRSCTVTINADGNFSGVSGSIWVN 515  
 RESULT 3  
 ID AMY\_BACST STANDARD; PRT; 549 AA.  
 AC P06279;  
 DT 01-JAN-1988 (REL. 06, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYS.  
 OS BACILLUS STEAROTHERMOPHILUS.

OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-39.  
 RX MEDLINE; 85234394.  
 RA NAKAJIMA R., IMANAKA T., AIBA S.;  
 RL J. BACTERIOL. 163:401-406(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-DY5/PHI300;  
 RX MEDLINE; 86008166.  
 RA IHARA H., SASAKI T., TSUBOI A., YAMAGATA H., TSUKAGOSHI N., UDAKA S.;  
 RL J. BIOCHEM. 98:95-103(1985).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-N2-3;  
 RX MEDLINE; 86195857.  
 RA GRAY G.L., MAINZER S.E., REY M.W., LAMSA M.H., KINDLE K.L.,  
 RA CARMONA C., REQUADT C.;  
 RL J. BACTERIOL. 166:635-643(1986).  
 RN [4]  
 RP SEQUENCE OF 1-122 FROM N.A., AND SEQUENCE OF 35-48.  
 RC STRAIN-DY-5;  
 RX MEDLINE; 86059211.  
 RA TSUKAGOSHI N., IRIANI S., SASAKI T., TAKEMURA T., IHARA H.,  
 RA IDOTA Y., YAMAGATA H., UDAKA S.;  
 RL J. BACTERIOL. 164:1182-1187(1985).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; M11450; G142505; -;  
 DR EMBL; X02769; G580825; -;  
 DR EMBL; M13255; G142513; -;  
 DR PIR; A00845; ALBSF.  
 DR PIR; A24549; A24549.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.  
 FT SIGNAL 1 34  
 FT CHAIN 35 549 ALPHA-AMYLASE.  
 FT ACT\_SITE 268 268 BY SIMILARITY.  
 FT ACT\_SITE 272 272 BY SIMILARITY.  
 FT ACT\_SITE 365 365 BY SIMILARITY.  
 FT CONFLICT 13 13 M -> V (IN REF. 3).  
 FT CONFLICT 19 19 L -> W (IN REF. 3).  
 FT CONFLICT 23 23 L -> S (IN REF. 2 AND 3).  
 FT CONFLICT 31 31 P -> H (IN REF. 2 AND 4).  
 FT CONFLICT 107 107 A -> T (IN REF. 2 AND 3).  
 FT CONFLICT 179 179 P -> N (IN REF. 3).  
 FT CONFLICT 251 251 S -> N (IN REF. 2 AND 3).  
 FT CONFLICT 284 284 D -> Y (IN REF. 2 AND 3).  
 FT CONFLICT 312 312 M -> T (IN REF. 2 AND 3).  
 FT CONFLICT 338 338 T -> A (IN REF. 2 AND 3).  
 FT CONFLICT 342 342 R -> S (IN REF. 3).  
 FT CONFLICT 346 346 T -> N (IN REF. 3).  
 FT CONFLICT 376 376 V -> C (IN REF. 2 AND 3).  
 FT CONFLICT 526 527 WS -> RP (IN REF. 2).  
 FT CONFLICT 527 527 S -> P (IN REF. 2).  
 FT CONFLICT 535 535 D -> G (IN REF. 2 AND 3).  
 SQ SEQUENCE 549 AA; 62670 MW; 4C7BEAD6 CRC32;  
 Query Match 68.3%; Score 2645; DB 1; Length 549;  
 Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
 Matches 339; Conservative 81; Mismatches 80; Indels 5; Gaps 5;  
 Db 15 LLAFLT-ALLFCPTGQAKAA-AFENGTMQYFEMWLPDDGLTWKTVANEANLSSIGI 72  
 Qy 11 LUTLLAVAVLPYPTEPAQAHHNTGNTGMQYFEWHLPDNGNHNWRDLDAANLKSIGI 70  
 Db 73 TALWLPAYKTSRSDVGYGYVLDLGEFNQKGAVRTKYGTKAQYLO-AIQAAHAAGMQ 131  
 Qy 71 TAVWIPPAWKQSNDVGYGYVLDLGEFNQKGAVRTKYGTQSQ-LOGAVTSLKNNGIQ 129  
 Db 132 VYADVDFHKGADGTEWDAVEVNPDSRQBEISCTYIOQAWTKFDFPGRGNTYSSFKWR 191  
 FT SIGNAL 1 31  
 FT CHAIN 32 514 ALPHA-AMYLASE.

Qy 130 VYGVVNHKGGADGTEWDAVEVNPDSRQBEISCTYIOQAWTKFDFPGRGNTYSSFKWR 189  
 Db 192 WYHFDGYDWDSEKRL-SRIYKFRGIGKAWDEVDTEGNTDYLMYADLMDHDPVVTTELK 250  
 Qy 190 WYHFDGTDWQSRQLQNKIKYKFRGTGRKAWDEVDIENGNDYLMYADLMDHDPVINELR 249  
 Db 251 SWGKVVNTNIDGFLDAVKHKKFSFPDWLSVRSQTKPLFTVGEYSYDINKLHNY 310  
 Qy 250 NWGYYTNTLNLGDFRIDAVKHKYSYTRDWLTHVRNTTGKPMFAVAFWKNDLAAIENY 309  
 Db 311 IMKTNGTMSLFDAPLHNKFTASKSGGTEDMRTLMTLTKMDQPTLAVTFVDNHDTEPGQ 370  
 Qy 310 LNKTSWNHVSFEDVPLHNLNNSNGGTFDMRNILSVVQKHPIHAVTFVDNHDSPQPE 369  
 Db 371 ALQSWDVPWFKPLAYALITRQEGYPCVFYGYGYQYINIPISLKSIDPLLIARRDYAY 430  
 Qy 370 ALESFVQSWFKPLAYALITRQEGYPSVFYGYGYGITHGVPSMKSIDPLLQARQYAY 429  
 Db 431 GTQHDYLDHSDIIGWTRGVTEKPGSLAALITDGPQSGKMWYGVKGHAGKVFYDLTGNR 490  
 Qy 430 GTQHDYFDHDDIIGWTRGDSHSPNSGLATIMSDPGGNKMWYGVKGHAGQVWRDITGNR 489  
 Db 491 SDVTINSDGWEKFGVNGSVVWV 515  
 Qy 490 SGTVTINADGNGNETVNGGAVVWV 514  
 RESULT 4  
 ID AMY BACAM STANDARD; PRT; 514 AA.  
 AC P00692;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 OS BACILLUS AMYLOLIQUEFACIENS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-IH;  
 RX MEDLINE; 83108808.  
 RA TAKKINEN K., PETERSSON R.F., KALKKINEN N., PALVA I., SODERLUND H.,  
 RA KAARTAINEN L.;  
 RL J. BIOL. CHEM. 258:1007-1013(1983).  
 RN [2]  
 RP SEQUENCE OF 32-222.  
 RX MEDLINE; 80241725.  
 RA CHUNG H.S., FRIEDBERG F.;  
 RL BIOCHEM. J. 185:387-395(1980).  
 RN [3]  
 RP SEQUENCE OF 1-96 FROM N.A.  
 RX MEDLINE; 82051296.  
 RA PALVA I., PETERSSON R.F., KALKKINEN N., LEHTOVAARA P., SARVAS M.,  
 RA SODERLUND H., TAKKINEN K., KAARTAINEN L.;  
 RL GENE 15:43-51(1981).  
 RN [4]  
 RP SEQUENCE OF 1-39 FROM N.A.  
 RX MEDLINE; 88137952.  
 RA RUOHONEN L., HACKMAN P., LEHTOVAARA P., KNOWLES J.K.C., KARAENEN S.;  
 RL GENE 59:161-170(1987).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; J01542; G142429; -;  
 DR EMBL; V00092; G39298; -;  
 DR EMBL; A20154; G580683; -;  
 DR EMBL; M18424; G142431; -;  
 DR PIR; A00843; ALBSN.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.  
 FT SIGNAL 1 31  
 FT CHAIN 32 514 ALPHA-AMYLASE.

US-08-952-741-2.ISP

Wed Sep 16 09:49:09 1998

J. GEN. MICROBIOL. 138:1051-1065(1992).  
 [3]  
 RN SEQUENCE OF 476-494 FROM N.A.  
 RP RAHA M., KIHARA M., KAWAGISHI I., MACNAB R.M.;  
 RA SUBMITTED (MAR-1993) TO ENDOGENOUS/DBJ DATA BANKS.  
 RL CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL: L01643; G154045; -  
 DR EMBL: M85241; G153989; -  
 DR EMBL: L13280; G295194; -  
 DR PIR: B45738; B45738.  
 DR STYGENE; SG10011; AMYA.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM.  
 FT ACT\_SITE 235 BY SIMILARITY.  
 FT ACT\_SITE 239 BY SIMILARITY.  
 FT ACT\_SITE 332 BY SIMILARITY.  
 FT ACT\_SITE 332 BY SIMILARITY.  
 SQ SEQUENCE 494 AA; 56496 MW; 2F27D4B3 CRC32;  
 Query Match 36.2%; Score 1403; DB 1; Length 494;  
 Best Local Similarity 42.5%; Pred. No. 1.24e-289;  
 Matches 209; Conservative 106; Mismatches 159; Indels 18; Gaps 15;  
 Db 3 NPTLQYFHWYDGGKLSWELAEADGLNDGINMVLPPACKGASGGYSGYDYDLF 62  
 QY 37 NGTMQYFEWHLNDGNHNRRLDDAANLKSIGITAVWIPPAWKGTSQNDVGYGAYDLY 95  
 Db 63 DLGFEQKGTIATKYDKRQLLTAIDALKNNIAVLDDVNVNKKGADEKERIVQRVQ 122  
 QY 96 DLGFEQKGTIATKYDKRQLLTAIDALKNNIAVLDDVNVNKKGADEKERIVQRVQ 155  
 Db 123 DRTQIDNNIECEGTRVTFPARGQSNFTWD-YHCFSGIDHINPD-EDGIFKIYND 180  
 QY 156 SNRQESGSEYTIETAWTKDFDGRGTHSNFKRWYH-EDGTDMDQSLQNKIYF-RG 213  
 Db 181 YTGQWQDQVDEMGDFDYLMDGENIDFRNHATEEIKYKAWVMEQTHCDGRDQVAKHI 240  
 QY 214 TGRWDEVDIENGNDYLMYADIDMDHPEVINELRNMGVWYTNLNDGFRIDAVKHI 272  
 Db 241 PAMFVKEWIEHVQAVKPLFIVAEVSWHSDVKLOTYYIDVDGKTMDFADPQMKFHAS 300  
 QY 273 KYSTRDMLTHVNTTQKPMFAVAFKNDLAAIENLNKTSWNHNSVDFVPLHYLNAS 332  
 Db 301 ROGAEDMRHIFTGLVEADPFHATLVVAHDTQPOLEAPVPEWPKPLAYALLILREN 360  
 QY 333 NSGGYFDMRNLGNSVQKHPHATVFDVNDHSDQGEALSFVQSWFKPLAYALLILREQ 392  
 Db 361 GVPVSFYPDLYCASVEDSGENGETCRVDMVPVINDLRLILARQFAHGTQTFEHPNCI 420  
 QY 393 GYPSVFYGD-Y--YGIPT-HG-V-P-SMK--SKIDPLQARQTYAYGTQHDYFDHDDII 442  
 Db 421 AFSRSGTEENPGC-VV-VLSNGDDGKTLILLGDYANKTWRDFSGNRDEYVVTNDQGEAT 478  
 QY 443 GWRREGDSSHPNSGLATIMSDPGGNKMYVCKHAKAGVWRDITGNSRGTVINADGWN 502  
 Db 479 FFCNAGSVVWV 490  
 QY 503 FTNGGAVSVWV 514  
 RESULT 6 STANDARD; PRT; 495 AA.  
 ID AMY2.ECOLI  
 AC P26613; P78072;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CYTOPLASMIC ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYA.  
 OS ESCHERICHIA COLI.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 FT ACT\_SITE 262 BY SIMILARITY.  
 FT ACT\_SITE 266 BY SIMILARITY.  
 FT ACT\_SITE 359 BY SIMILARITY.  
 FT ACT\_SITE 54 L -> I (IN REF. 2).  
 FT ACT\_SITE 64 L -> L (IN REF. 2).  
 FT ACT\_SITE 79 S -> D (IN REF. 2).  
 FT ACT\_SITE 84 G -> S (IN REF. 2).  
 FT ACT\_SITE 84 G -> S (IN REF. 2).  
 SQ SEQUENCE 514 AA; 58403 MW; A37712F3 CRC32;  
 Query Match 66.1%; Score 2559; DB 1; Length 514;  
 Best Local Similarity 64.0%; Pred. No. 0.00e+00;  
 Matches 330; Conservative 90; Mismatches 89; Indels 7; Gaps 5;  
 Db 1 MIQKRTVFRVLVCTLLF-VSLPI-TKTSAYNGTLMQYFETVNDQKHRLQND 57  
 QY 2 KLNRIISVLLTLLAVLVFPYTPAQAHNGTNGTMQYFEWHLNDGNHNRRLDD 61  
 Db 58 AEHLSDIGITAVWIPPAWKGTSQNDVGYGAYDLYDGEFQKGTIATKYDKRQLLTAID 117  
 QY 62 AANLKSIGITAVWIPPAWKGTSQNDVGYGAYDLYDGEFQKGTIATKYDKRQLLTAID 121  
 Db 118 SLHSRVQYGVNHNKAGADATEDVAVENPANRNOETSEYQIKAWTDFRPGRN 177  
 QY 122 SLKNGIQYGVNHNKAGADATEDVAVENPANRNOETSEYQIKAWTDFRPGRN 181  
 Db 178 TYSDFKWHYHFDGADWDESKRI-SRIKPEGKANDWVSSSENGNDYLYADVDYDH 236  
 QY 182 THSNFKRWYHFDGADWDESKRI-SRIKPEGKANDWVSSSENGNDYLYADVDYDH 241  
 Db 237 PDVVAETKKGWYANLSDGFRIDAAKHIFSLRDWQVQVQATGKEMFTVAEYWN 296  
 QY 242 PEVINELRNMGVWYTNLNDGFRIDAVKHIKYSYTRDMLTHVNTTQKPMFAVAFWKN 301  
 Db 297 NAGLENYLNKTSWNHNSVDFVPLHYLNASNSGGYFDMRNLGNSVQKHPHATVFD 356  
 QY 302 DLAAIENLNKTSWNHNSVDFVPLHYLNASNSGGYFDMRNLGNSVQKHPHATVFD 361  
 Db 357 NHDTPQGSLESTVQWFKFLAYAFILTRSGYPOVFGDMYKGTSPKIPSLKONIE 416  
 QY 362 NHDTPQGSLESTVQWFKFLAYAFILTRSGYPOVFGDMYKGTSPKIPSLKONIE 418  
 Db 417 PILKAREYAGPOHDIIDHPDVGWTRGDSAAKSLAALITDGPQSKRMVAGLNA 476  
 QY 419 PLLQARQTYAYGTQHDYFDHDDIIIGWTRGDSHNSGLATIMSDPGGNKMYVCKHAK 478  
 Db 477 GETWYDITGNSRDTYKIGSDGSEFHHVNDGVSIIYV 512  
 QY 479 GQVWRDITGNSRGTVINADGNGFTVNGGAVSVWV 514  
 RESULT 5 STANDARD; PRT; 494 AA.  
 ID AMY2.SALTY  
 AC P26613;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE CYTOPLASMIC ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYA.  
 OS SALMONELLA TYPHIMURIUM.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SJW1103;  
 RX MEDLINE; 93015717.  
 RA RAHA M., KAWAGISHI I., MUELLER V., KIHARA M., MACNAB R.M.;  
 RL J. BACTERIOL. 174:6644-6652(1992).  
 RN [2]  
 RP SEQUENCE OF 1-6 FROM N.A.  
 RC STRAIN-SJW1103;  
 RX MEDLINE; 92407478.  
 RA KAWAGISHI I., MUELLER V., WILLIAMS A.W., IRIKURA V.M., MACNAB R.M.;

```

OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JALL;
RX MEDLINE; 93015717.
RA RAHA M., KAWAGISHI I., MUELLER V., KIHARA M., MACNAB R.M.;
RL J. BACTERIOL. 174:6644-6652(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RC BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE; 97251358.
RA ITOH T., ATBA H., BABU T., FUJITA K., HAYASHI K., INADA T.,
RA ISONO K., KASAI H., KIMURA S., KITAKAWA M., KITAGAWA M.,
RA MAKINO K., MIKI T., MIZOGUCHI K., MORI H., MORI T., MOTOMURA K.,
RA NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,
RA SAITO N., SAMPEI G., SERI Y., SIVASUNDARAM S., TAGAMI H.,
RA TAKEDA J., TAKEMOTO K., WADA C., YAMAMOTO Y., HORIUCHI T.;
RL DNA RES. 3:379-392(1996).
RN [4]
RP SEQUENCE OF 1-5 FROM N.A.
RX STRAIN=JALL;
RX MEDLINE; 92407478.
RA KAWAGISHI I., MUELLER V., WILLIAMS A.W., IRIKURA V.M., MACNAB R.M.;
RL J. GEN. MICROBIOL. 138:1051-1065(1992).
RN [5]
RP SEQUENCE OF 475-495 FROM N.A.
RX STRAIN=JALL;
RX MEDLINE; 93381452.
RA RAHA M., KIHARA M., KAWAGISHI I., MACNAB R.M.;
RL J. GEN. MICROBIOL. 139:1401-1407(1993).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -!- SUBCELLULAR LOCATION: CYTOSOLASMIC.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLAZE FAMILY.
DR EMBL; L01642; G146023; -
DR EMBL; AE000285; G178236; -.
DR EMBL; D90833; G1736595; -.
DR DR EMBL; M85240; E56260; -.
DR EMBL; L13279; G290412; -.
DR PIR; A45738; A45738.
DR ECOGENE; EG11387; AMYA.
KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM.
FT ACT_SITE 235 235 BY SIMILARITY.
FT ACT_SITE 239 239 BY SIMILARITY.
FT ACT_SITE 332 332 BY SIMILARITY.
FT CONFLICT 19 20 KL -> SS (IN REF. 1).
FT CONFLICT 109 109 A -> V (IN REF. 1).
FT CONFLICT 149 149 Q -> E (IN REF. 1).
FT CONFLICT 234 234 L^--> I (IN REF. 1).
SQ SEQUENCE 495 AA; 56639 MW; 3A9A2183 CRC32;

Query Match 35.1%; Score 1360; DB 1; Length 495;
Best Local Similarity 41.7%; Pred.No.1.98e-279;
Matches 205; Conservative 101; Mismatches 168; Indels 18; Gaps 1

Db 3 NPITLLQCFHWYEGKLPEALAERADGFGINNVLPPAYKGASGYSGVDSYDLF 62
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 37 NGTMQYFEWHLPDNGHNRLRDDRDAANLKSGITAVLPAPWKGTSON-DVGYGAYDY 95
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

Db 63 DLGEFTQKGSIPTKYGDKAQLAAIALDKRRNDIAVLDDVVVNHKMGADKEAIRVQRVNA 122
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 96 DLGEFNQKGVTRTKYSTRSLOQAVTSLNKNGIQVYGDVVMNHKGGADGCTEMVNAVEYNR 155
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

Db 123 DDRTO-IDEEIIICEGWTRYTPPARAGQVSQIFWDFKCFSGIDHDHENPD-EDGIKFIVND 180
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 156 SNRNQGISGETT-EAWTKFDFFGRGNTSNFKRWYHFDDGTDWDQSRLQKNKYKF-RG 213
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

```



RESULT	12	STANDARD;	PRT;	718 AA.
ID	CDGT_BACSS			
AC	P31747;			
DT	01-JUL-1993	(REL. 26, CREATED)		
DT	01-JUL-1993	(REL. 26, LAST SEQUENCE UPDATE)		
DT	01-JUL-1993	(REL. 26, LAST ANNOTATION UPDATE)		
DE	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR			
DE	(CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE)			

ID	CDGT_BACSS	STANDARD:	P31747:	718 AA
AC				
DT	01-JUL-1993	(REL. 26, CREATED)		
DT	01-JUL-1993	(REL. 26, LAST SEQUENCE UPDATE)		
DT	01-JUL-1993	(REL. 26, LAST ANNOTATION UPDATE)		
DE	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR			
DE	(CYCLOMALTODEXTRIN-GLYCOSYLTRANSFERASE) CGCTASE			

GN CGT.  
OS BACILLUS SP. (STRAIN 6.6.3).  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA AKHMETZJANOV A.A.;  
RL SUBMITTED (MAY-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION  
CC OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.  
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-  
CC TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH,  
CC AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES,  
CC INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE  
CC FOR CYCLIZING THE MALTOLOGOSACCHARIDE PRODUCED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
DR EMBL: X66106; G39839; -.  
DR PIR: S21532; ALBSG6.  
DR HSSP: P30920; ICGT.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL.  
FT SIGNAL 1 34 POTENTIAL.  
FT CHAIN 35 718  
FT DOMAIN 35 172  
FT DOMAIN 173 236  
FT DOMAIN 237 440  
FT DOMAIN 441 528  
FT DOMAIN 529 614  
FT DOMAIN 615 718  
FT DISULFID 77 84  
FT ACT\_SITE 263 263 BY SIMILARITY.  
FT ACT\_SITE 291 291 BY SIMILARITY.  
FT ACT\_SITE 362 362 BY SIMILARITY.  
SQ SEQUENCE 718 AA; 73784DA0 CRC32;  
  
Query Match 4.8%; Score 189; DB 1; Length 718;  
Best Local Similarity 24.2%; Pred. No. 4.62e-15;  
Matches 52; Conservative 69; Mismatches 71; Indels 23; Gaps 21;  
  
Db 220 LENGIYKNL-YDLADFNHNNATIDKYFKAIKWL-D-MGVGIRVDAVKHIALGWQSW 276  
QY 224 IENGNDYLMYADIMDHPEV-INE-LRNW-GWYNTNLNLDGFRIDAVKHKYSYTRDW 276  
Db 277 MSSI-IVH-KPVFTGEGWFLGSAADANTDFANKSGMSLLDFRNSAVR-NVFRDNTSN 280  
QY 281 LTHVRNTTKPMFAVAE-FWKNDLAIEN--YLNKTSWNH-SV-FDVPPLHYLNASNSG 333  
Db 334 MY-ALDSMINSTATDYNQVDFIDNMDRFT-SAVNNRLEQALAFILT-SR--G 388  
QY 336 GFYDMRNILGVSQKPIHA-VTFVNDHDSQGEALSFQSWF-KPLAYAILITREQ 393  
Db 389 VPAIYIGTEQY-LTGNGDDPDNRKMPFSKSTTAF 422  
QY 394 YPSVFYG-DYIGTHGVPSMKSIDPLQARQTY 427  
  
RESULT 13  
ID CCGU\_BACCI STANDARD; PRT; 713 AA.  
AC P43379;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE CYCLOMALTOEXTRIN GLUCANOTRANSFERASE (CYCLOMALTOEXTRIN)  
DE (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).  
OS BACILLUS CIRCULANS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RC STRAIN-251;  
RX MEDLINE; 94149761.  
RA LAWSON C.L., VAN MONTFORT R., STROKOPYTOV B., ROZEBOOM H.J.,  
RA KALK K.H., DE VRIES G.E., PENNINGA D., DIJKHUIZEN L., DIJKSTRA B.W.,

RL J. MOL. BIOL. 236:590-600(1994).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC STRAIN-251;  
RX MEDLINE; 96094317.  
RA KNEGTEL R.M.A., STROKOPYTOV B., PENNINGA D., FABER O.G.,  
RA ROZEBOOM H.J., KALK K.H., DIJKSTRA B.W., DIJKHUIZEN L.,  
RL J. BIOL. CHEM. 270:29256-29264(1995).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE; 97115811.  
RA PENNINGA D., VAN DER VEEN B.A., KNEGTEL R.M., VAN HIJUM S.A.,  
RA ROZEBOOM H.J., KALK K.H., DIJKSTRA B.W., DIJKHUIZEN L.,  
RL J. BIOL. CHEM. 271:32777-32784(1996).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH INHIBITOR.  
RC STRAIN-251;  
RX MEDLINE; 96264806.  
RA STROKOPYTOV B., KNEGTEL R.M.A., PENNINGA D., ROZEBOOM H.J., KALK K.H.,  
RA DIJKHUIZEN L., DIJKSTRA B.W.;  
RL BIOCHEMISTRY 35:4241-4249(1996).  
CC -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION  
CC OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.  
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- DOMAIN: MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-  
CC TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH, AND  
CC THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES,  
CC INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE  
CC FOR CYCLIZING THE MALTOLOGOSACCHARIDE PRODUCED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
DR EMBL: X78145; G510492; -.  
DR PDB: 1CXG; 20-DEC-94.  
DR PDB: 1CDG; 08-MAR-95.  
DR PDB: 1CGV; 27-FEB-95.  
DR PDB: 1CGW; 27-FEB-95.  
DR PDB: 1CGX; 07-FEB-95.  
DR PDB: 1CGY; 07-FEB-95.  
DR PDB: 1CXE; 15-FEB-95.  
DR PDB: 1CAF; 15-DEC-95.  
DR PDB: 1CXH; 15-DEC-95.  
DR PDB: 1CXI; 15-DEC-95.  
DR PDB: 1DIJ; 21-APR-97.  
DR PDB: 1TCM; 21-APR-97.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL; 3D-STRUCTURE.  
FT SIGNAL 1 25  
FT CHAIN 26 713  
FT DOMAIN 28 165 CYCLOMALTOEXTRIN GLUCANOTRANSFERASE.  
FT DOMAIN 166 229 A1.  
FT DOMAIN 230 433 B.  
FT DOMAIN 434 522 A2.  
FT DOMAIN 523 609 C.  
FT DOMAIN 610 713 D.  
FT DISULFID 70 77 E.  
FT ACT\_SITE 256 256 NUCLEOPHILE.  
FT ACT\_SITE 284 284 PROTON DONOR.  
FT ACT\_SITE 355 355 SUBSTRATE-BINDING.  
SQ SEQUENCE 713 AA; 77309 MW; 3DC02138 CRC32;  
  
Query Match 4.8%; Score 185; DB 1; Length 713;  
Best Local Similarity 25.0%; Pred. No. 2.48e-14;  
Matches 53; Conservative 64; Mismatches 76; Indels 19; Gaps 18;  
  
Db 214 ENGIYKNL-YDLADFNHNNSTVDYVKDAIKWL-D-LGIDGIRMDAVKHPFGQKSF 270  
QY 225 ENGNIDYLMYADIMDHPE-VINE-LRNW-GWYNTNLNLDGFRIDAVKHKYSYTRDW 270  
Db 271 AAV-NNY-KPVFTGEGWFLGVSNEVSPENHAFANESGMSLLDFRFAQKRVQFRD-NTDNM 281  
QY 282 THVRNTTKPMFAVAE-FWKNDLAIEN-YLNKTSWNHVSFVDFVPLHY-NLYNASNGGY 327  
Db 328 YGLKAMLEGSADYQVDDQVTFIDNHDHMERFASNNRRKLEQALAFILT-SR--GVPA 384



QY	338	FDNRNLINGSVQKPHIA-VTFVDNHDSPGSALESFVQSMFKPLAYALILTREQGPS	396
QY	385	IYYTEQY-MSGGTPDPNRRARIPSFSTSTAY	415
QY	397	VFYG-DYGIPTHGVPMSKIDPLQARQTY	427
RESULT	14		
ID	COGT_BACCI	STANDARD;	PRT; 718 AA.
AC	P30920;		
DT	01-JUL-1993	(REL. 26, CREATED)	
DT	01-JUL-1993	(REL. 26, LAST SEQUENCE UPDATE)	
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)	
DE	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE	PRECURSOR (EC 2.4.1.19)	
DE	(CYCLODEXTRIN-GLYCOSYLTRANSFERASE)	(CGTASE).	
OS	BACILLUS CIRCULANS.		
OC	PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=8;		
RX	MEDLINE; 91103970.		
RX	NITSCHKE L., HEEGER K., BENDER H., SCHULZ G.E.;		
RL	APPL. MICROBIOL. BIOTECHNOL. 33:542-546(1990).		
RN	[2]		
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).		
RC	STRAIN=8;		
RX	MEDLINE; 91171298.		
RA	KLEIN C., SCHULZ G.E.;		
RL	J. MOL. BIOL. 217:737-750(1991).		
RN	[3]		
RP	X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).		
RA	MEDLINE; 90064533.		
RX	HOFMANN B.E., BENDER H., SCHULZ G.E.;		
RL	J. MOL. BIOL. 209:793-800(1989).		
CC	-!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.		
CC	-!- COFACTOR: BINDS TWO CALCIUM IONS.		
CC	-!- SUBUNIT: MONOMER.		
CC	-!- CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE MALTOOLIGOSACCHARIDE PRODUCED.		
CC	-!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.		
DR	EMBL; X68326; G39420; -		
DR	PIR; S23674; ALBSCG.		
DR	PDB; 1CGT; 31-JAN-94.		
DR	PDB; 1CGJ; 31-JAN-94.		
KW	TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL; 3D-STRUCTURE.		
FT	SIGNAL	1 34	
FT	CHAIN	35 718	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
FT	DOMAIN	35 172	A1.
FT	DOMAIN	173 236	B.
FT	DOMAIN	237 440	A2.
FT	DOMAIN	441 528	C.
FT	DOMAIN	529 614	D.
FT	DOMAIN	615 718	E.
FT	DISULFID	77 84	
FT	ACT_SITE	258 258	BY SIMILARITY.
FT	ACT_SITE	291 291	BY SIMILARITY.
FT	ACT_SITE	362 362	BY SIMILARITY.
FT	TURN	40 41	
FT	TURN	43 44	
FT	TURN	47 48	
FT	STRAND	51 53	
FT	HELIX	56 59	
FT	TURN	71 72	
FT	TURN	73 73	
FT	TURN	75 76	
FT	TURN	80 81	
FT	TURN	83 83	



\*\*\*\*\*  
WILEY  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:19:35 1998; MasPar time 31.41 Seconds  
691.851 Million cell updates/sec

Tabular output not generated.

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHNRIISVLLTLLAVAV.....ADGWGFTVNGGAVSVVVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sptrembl6  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_nhc 6:sp\_organelle 7:sp\_phase 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp\_vertebrate  
13:sp\_unclassified

Statistics: Mean 49.603; Variance 91.345; scale 0.543

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB	Pred. No.
1	2672	69.0 513 9 Q59222	ALPHA-AMYLASE (EC 3.2. 0.00e+00
2	2648	68.4 521 9 P71034	ALPHA-AMYLASE PRECURSOR 0.00e+00
3	2642	68.2 549 9 Q31193	ALPHA-AMYLASE 0.00e+00
4	2608	67.3 549 9 Q45519	ALPHA-AMYLASE PRECURSOR 0.00e+00
5	1912	49.4 423 9 Q03657	AMYLASE (EC 3.2.1.-) 0.00e+00
6	1682	43.4 485 9 Q53786	AMYLASE 0.00e+00
7	368	9.5 461 9 Q33476	ALPHA-AMYLASE PRECURSOR 3.77e-49
8	362	9.3 460 9 Q08452	ALPHA-AMYLASE (EC 3.2. 6.41e-48
9	172	4.4 504 9 Q60224	ALPHA-AMYLASE PRECURSOR 2.05e-11
10	150	3.9 482 9 Q60051	ALPHA-AMYLASE PRECURSOR 9.61e-08
11	148	3.8 423 8 Q42678	ALPHA-AMYLASE PRECURSOR 2.02e-07
12	135	3.5 407 8 Q41442	ALPHA-AMYLASE PRECURSOR 2.27e-05
13	135	3.5 427 8 Q03651	ALPHA-AMYLASE TYPE B P 2.27e-05
14	130	3.4 349 8 Q41441	ALPHA-AMYLASE (FRAGMEN 1.31e-04
15	132	3.4 421 8 Q42504	ALPHA-AMYLASE TYPE A. 6.53e-05
16	131	3.4 427 8 Q40015	ALPHA-AMYLASE 1. 9.26e-05
17	130	3.4 437 8 Q04965	ALPHA-AMYLASE (EC 3.2. 1.31e-04
18	131	3.4 438 8 Q40017	ALPHA-AMYLASE 2. 9.26e-05
19	128	3.3 429 8 Q40016	ALPHA-AMYLASE 1. 2.62e-04
20	127	3.3 499 1 Q00250	TAKA-AMYLASE A (TAA-G1 3.70e-04

21	129	3.3	1590	9 Q59983	GLUCOSYLTRANSFERASE-I	1.86e-04
22	129	3.3	1590	9 Q55263	GLUCOSYLTRANSFERASE GT	1.86e-04
23	125	3.2	437	8 Q04964	ALPHA-AMYLASE (EC 3.2. 7.33e-04	
24	122	3.2	640	1 Q13296	ACID-STABLE ALPHA-AMYL	2.02e-03
25	120	3.1	439	8 Q41770	ALPHA-AMYLASE	3.93e-03
26	121	3.1	513	1 Q41154	ALPHA AMYLASE PRECURSO	2.82e-03
27	120	3.1	552	9 Q44528	PUTATIVE ALPHA-GLUCANO	3.93e-03
28	119	3.1	700	9 Q24781	ALPHA-AMYLASE PRECURSO	5.48e-03
29	116	3.0	441	9 Q60035	4-ALPHA-GLUCANOTRANSFE	1.46e-02
30	116	3.0	475	9 Q56791	ALPHA-AMYLASE PRECURSO	1.46e-02
31	115	3.0	570	1 Q01117	ALPHA-AMYLASE	2.02e-02
32	115	3.0	590	9 Q45490	BSMA.	2.02e-02
33	118	3.0	725	9 Q59239	CYCLODEXTRIN GLUCANOTR	7.62e-03
34	114	2.9	437	8 Q40018	ALPHA-AMYLASE 2.	2.79e-02
35	111	2.9	466	9 Q52414	ALPHA-AMYLASE PRECURSO	7.26e-02
36	113	2.9	614	9 Q52516	MALTOPEPTASE FORMING	3.85e-02
37	113	2.9	652	9 Q59242	1,4-ALPHA-GLUCAN BRANC	3.85e-02
38	113	2.9	1290	9 Q48756	DEXTRANSUCRASE.	3.85e-02
39	114	2.9	1449	9 Q55264	GLUCOSYLTRANSFERASE PR	2.79e-02
40	107	2.8	381	9 Q25026	SIGNAL-TRANSDUCING PRO	2.53e-01
41	108	2.8	421	9 Q59243	TRUNCATED 1,4-ALPHA-GL	1.86e-01
42	108	2.8	491	1 Q13996	AMYLASE PRECURSOR	1.86e-01
43	107	2.8	507	1 Q08806	ALPHA-AMYLASE 2 (EC 3.	2.53e-01
44	107	2.8	1599	9 Q00599	GLUCOSYLTRANSFERASE S	2.53e-01
45	105	2.7	394	11 Q04537	GENOME, PARTIAL SEQUEN	4.65e-01

ALIGNMENTS

RESULT 1  
ID Q59222 PRELIMINARY; PRT; 613 AA.  
AC Q59222;  
DT 01-NOV-1996 (TREMUREL. 01, CREATED)  
DT 01-NOV-1996 (TREMUREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMUREL. 01, LAST ANNOTATION UPDATE)  
DE ALPHA-AMYLASE (EC 3.2.2.1) (PURINE NUCLEOSIDASE).  
GN AMY.  
OS BACILLUS SP.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TS-23;  
RA LIN L.L., CHU W.S., HSU W.H.;  
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: A N-D-RIBOSYLPURINE + H(2)O -  
CC A PURINE + D-RIBOSE.  
DR EMBL; U22045; G72279; -.  
KW HYDROLASE; GLYCOSIDASE.  
SQ SEQUENCE 613 AA; 69537 MW; DC66587C CRC32;

Query Match				69.0%; Score 2672; DB 9; Length 613;
Best Local Similarity				68.5%; Pred. No. 0.00e+00;
Matches				342; Conservative 78; Mismatches 76; Indels 3; Gaps 2;
Db	17	LVISFFTPFST--AQANTAPINETMMQYFEWDLNDGTLTKVKNAAANLSLGTALWL	74	
Qy	16	LAVAVLPYPTPAQAHHNGTNGTMMQYFEWHLNDGNHNRDLRDDAANLSKGTAWI	75	
Db	75	PPAYKGTQSDVGVDYLDLGEFNQKGTIRTKYGTQYQIAQAAKAGMQYADV	134	
Qy	76	PPAWKGTQSDVGVDYLDLGEFNQKGTIRTKYGTQYQIAQAAKAGMQYADV	135	
Db	135	FNHKGAGDGTFEVDVPSNRNQTSGTQIQIANTKDFPGRGNTYSSPKRWYHFDG	194	
Qy	136	MNHKGAGDGTFEVDVPSNRNQTSGTQIQIANTKDFPGRGNTYSSPKRWYHFDG	195	
Db	195	TDWDSRLK-NRIYKFRSTGKAWDEVDENGVDYLMFADLMDHPVETLKNWGTW	253	
Qy	196	TDWDSRLK-NRIYKFRSTGKAWDEVDENGVDYLMFADLMDHPVETLKNWGTW	255	
Db	254	VNTNIDGFRDVAHKIKYSFFPDMLTYYVRNQTGKNLFVAGFEWSDYVKNLHNYTKNG	313	
Qy	256	TNTLNLDGFRDVAHKIKYSFFPDMLTYYVRNQTGKNLFVAGFEWSDYVKNLHNYTKNG	315	

Db 314 SMSLFDAPLHNNYTTASKSGGYEDMYRLNNTLMKDQPSLAVTLVDNHDTPQGSLSQSV 373  
QY 316 NHSVDFVPLHYNLNASGSGYEDMRNINLGSVVQKHPHIAVTFVDNHDSPQCEALESFV 375

Db 374 EPWFKPLAYAFILITROEGYPCVFGYDYGTPKYNIPGLSKSIDPLLIARDYAYGTFQDY 433  
QY 376 QSWFKPLAYALILITROGYPVFGYDYGTHGVFSMKSIDPLLOARQTYAGTQHDY 435

Db 434 IDHQDIIGWTREGIDTKPNPSGLAALITDGPGGSKMYVGGKHKAGKYFYDLTGNRSPTVTI 493  
QY 436 FDHDDIIGWTREGDSSHPNSGLATIMSDGPGGNKMYVGGKHKAGQVWRDITGNRSCTVTI 495

Db 494 NADGWGEFKYNGSVSIWV 512  
QY 496 NADGWGNTVNGGAVSVWV 514

RESULT 2 PRELIMINARY; PRT; 521 AA.  
ID P71034;  
AC P71034;  
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE PRECURSOR.  
OS BACILLUS SP.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MK 716;  
RA SIDHU G.S., CHAKRABARTI T.;  
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U75445; G1667474; -.  
KW SIGNAL.  
FT CHAIN 35 521  
SQ SEQUENCE 521 AA; 59311 MW; 55960B19 CRC32;

Query Match 68.4%; Score 2648; DB 9; Length 521;  
Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
Matches 339; Conservative 80; Mismatches 81; Indels 5; Gaps 5;

Db 15 LLAFLITAS-LFCPTGQPAKAA-APFNGTMMQYFEWLPDDGTLTWTKVANEANLSSLSGI 72  
QY 11 LTLTLLAVAVLFPYMTPEQAHHNGTNGTMMQYFEWHLPDNGNHNRLRODAANLKSIGI 70

Db 73 TALMLPPAYKGTSDVGYDYDLGFEFNQKAVRTKYGTAKQYLO-AIOAAHAAGMQ 131  
QY 71 TAVWIIPPAWKGTSQNDVGAYDYLDFEFGNKGVTYKGTGRSQ-LOGAVTSLKNGIQ 129

Db 132 VYADVVDHKGAGDGTWVDVAVENPSDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 71 TAVWIIPPAWKGTSQNDVGAYDYLDFEFGNKGVTYKGTGRSQ-LOGAVTSLKNGIQ 129

Db 132 VYADVVDHKGAGDGTWVDVAVENPSDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 130 VYGVVMMHKGAGDGTWVNAVEVNRNRNOEISGEYTIKFTKDFPGRNTHSNFKWR 189

Db 192 WYHFGVDWDESRKL-SRIYKFRGIGKAWDEVDTEGNYDYLMDADLMDHPEVVTTELK 250  
QY 190 WYHFGDWDQSRQONKIYKFRGTGKAWDEVDTEGNYDYLMDADLMDHPEVINELR 249

Db 251 NMGKVVNTNIDGFRDLDAVKHIFKFFPDMLSVYRSOTGKPLFTVGEYWSYDINKLHNY 310  
QY 250 NMGVWYNTNLNDGFRIDAVKHISYTRDMLTHVRNTTGGKPMFAVEFKNDLAAIENY 309

Db 311 ITKNTGMSLFDAPLHNNYTTASKSGGAFDMRTLMTNTLMKDQPTLAVTFVDNHDTPGQ 370  
QY 310 LNKTSWNHVSFVDFPLHYNLNASGSGYFDMRNILGSVVQKHPHIAVTFVDNHDSPQGE 369

Db 371 ALQSWDVPWFKPLAYAFILITROEGYPCVFGYDYGIPQYNIPSLKSIDPLLIARDYAY 430  
QY 370 ALESEFVQSWFKPLAYALILITROGYPVFGYDYGTHGVFSMKSIDPLLOARQTYAY 429

Db 431 GTQHDYLDHSDIIGWTREGVTEKPSGLAALITDGPGGSKMYVGGKHKAGKYFYDLTGNR 490  
QY 430 GTQHDYFDHDDIIGWTREGDSSHPNSGLATIMSDGPGGNKMYVGGKHKAGQVWRDITGNR 489

Db 491 SDVTINSDGWGEFKYNGSVSVWV 515  
QY 490 SGTVTINADGWGNTVNGGAVSVWV 514

RESULT 4 PRELIMINARY; PRT; 549 AA.  
ID Q45519;  
AC Q45519;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE PRECURSOR.  
OS BACILLUS STEAROTHERMOPHILUS.

Db 491 SDVTINSDGWGEFKYNGSVSVWV 515  
QY 490 SGTVTINADGWGNTVNGGAVSVWV 514

RESULT 3 PRELIMINARY; PRT; 549 AA.  
ID Q31193;  
AC Q31193;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE.  
GN AMI.  
OS BACILLUS STEAROTHERMOPHILUS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 31195;  
RA DA SILVA A.C.R., FERNANDES E., PUEYO M.T.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF032864; G2642326; -.  
SQ SEQUENCE 549 AA; 62651 MW; D9051082 CRC32;

Query Match 68.2%; Score 2642; DB 9; Length 549;  
Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
Matches 339; Conservative 79; Mismatches 82; Indels 5; Gaps 5;

Db 15 LLAFLITAS-LFCPTGQPAKAA-APFNGTMMQYFEWLPDDGTLTWTKVANEANLSSLSGI 72  
QY 11 LTLTLLAVAVLFPYMTPEQAHHNGTNGTMMQYFEWHLPDNGNHNRLRODAANLKSIGI 70

Db 73 TALMLPPAYKGTSDVGYDYDLGFEFNQKAVRTKYGTAKQYLO-AIOAAHAAGMQ 131  
QY 71 TAVWIIPPAWKGTSQNDVGAYDYLDFEFGNKGVTYKGTGRSQ-LOGAVTSLKNGIQ 129

Db 132 VYADVVDHKGAGDGTWVDVAVENPSDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 130 VYGVVMMHKGAGDGTWVNAVEVNRNRNOEISGEYTIKFTKDFPGRNTHSNFKWR 189

Db 192 WYHFGVDWDESRKL-SRIYKFRGIGKAWDEVDTEGNYDYLMDADLMDHPEVVTTELK 250  
QY 190 WYHFGDWDQSRQONKIYKFRGTGKAWDEVDTEGNYDYLMDADLMDHPEVINELR 249

Db 251 NMGKVVNTNIDGFRDLDAVKHIFKFFPDMLSVYRSOTGKPLFTVGEYWSYDINKLHNY 310  
QY 250 NMGVWYNTNLNDGFRIDAVKHISYTRDMLTHVRNTTGGKPMFAVEFKNDLAAIENY 309

Db 311 ITKNTGMSLFDAPLHNNYTTASKSGGAFDMRTLMTNTLMKDQPTLAVTFVDNHDTPGQ 370  
QY 310 LNKTSWNHVSFVDFPLHYNLNASGSGYFDMRNILGSVVQKHPHIAVTFVDNHDSPQGE 369

Db 371 ALQSWDVPWFKPLAYAFILITROEGYPCVFGYDYGIPQYNIPSLKSIDPLLIARDYAY 430  
QY 370 ALESEFVQSWFKPLAYALILITROGYPVFGYDYGTHGVFSMKSIDPLLOARQTYAY 429

Db 431 GTQHDYLDHSDIIGWTREGVTEKPSGLAALITDGPGGSKMYVGGKHKAGKYFYDLTGNR 490  
QY 430 GTQHDYFDHDDIIGWTREGDSSHPNSGLATIMSDGPGGNKMYVGGKHKAGQVWRDITGNR 489

Db 491 SDVTINSDGWGEFKYNGSVSVWV 515  
QY 490 SGTVTINADGWGNTVNGGAVSVWV 514

RESULT 4 PRELIMINARY; PRT; 549 AA.  
ID Q45519;  
AC Q45519;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE PRECURSOR.  
OS BACILLUS STEAROTHERMOPHILUS.

Best Local Similarity 51.5%; Pred. No. 0.00e+00;

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 104

QY 36 TNGTMMQYFEWHL.PNDGNHNRRLRDAANLKSIGITAVVWIPPAWKGTSONDVGYGAYDLY 95  
 Db 62 DLGEFNQKGTVTKYGTAKADYLAISALKNNGIEPLADYILNKAADHTETFKVVEVAP 121  
 QY 96 DLGEFNQKGTVTKYGTAKADYLAISALKNNGIEPLADYILNKAADHTETFKVVEVAP 155  
 Db 122 EDRTKVLSOPPEIEGWTNTEFEVAIVPYNDFECHYHGTGTYD-VKTKGTGIFQIOQDN 180  
 QY 156 SNRNOEISGEYITAEWTXFDFFGRGNTHSNFKRWYHFDGTDQDSROLQNKIYKFRGTG 215  
 Db 181 KGANODLVGNGYDYLMDADLKHPKPEVKNKYDADWVEVTKGPFELDAIKHID 240  
 QY 216 KAW-DWE-VDIENGNDYLMADIDMDHPEVINELRNGVWTTNLDGFRIDAVKHK 273  
 Db 241 SFFMGNFTRDMKTKSGNDFYVGEFWNGDEKSNNDYLASTDYRFDLVDRHLQNLFEASK 300  
 QY 274 YSYTRDWTHTVRNTGPKMFAVEFKNDLAAIENLYNKTSMNHSVDFVPLHYNLYNASN 333  
 Db 301 AKETYDLRQIFEOTLVKNHPSAVTFVDNHDTRQOALESTIEWFKPAAYALILLRGTG 360  
 QY 334 SGGYFDMRNILNGSVVQKHPHNAVTFVDNHDTRQOALESTIEWFKPAAYALILLRGTG 393  
 Db 361 LPCIFYDYGISGQAFQESQTVIDKLIELRKNVAGQEMDYDOANCIQWTCIGDDDEH 420  
 QY 394 YPSVFGYDYGIPTH-GVPSMKSKIDPLQARQIYAYGTQHDYFDHDDHILGTREGDSSH 452  
 Db 421 P-TALAGLINNSKATSRMFKVGEKWACKLFTDALGNQAAHVQIDEGQYDGLVGEKSVSA 479  
 QY 453 PNLGATIMSDPGGKNWYVKGKAGQVWRDITGNRSGTITINADGWNFTVNGGAVSV 512  
 Db 480 WI 481  
 QY 513 WV 514

RESULT 7  
 ID O33476 PRELIMINARY; PRT; 461 AA.  
 AC O33476;  
 DT 01-JAN-1998 (TREMBLREL, 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL, 05, LAST SEQUENCE UPDATE)  
 DE ALPHA-AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN  
 GN APKA.  
 OS PYROCOCUS SP.  
 OC ARCHAEABACTERIA; CRENARCHAEOTA; THERMOPROTEALES; THERMOCOCCACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KOD1;  
 RA TACHIBANA Y., MENDEZ L., FUJIWARA S., TAKAGI M., IWANAKA T.;  
 RL J. FERMENT. BIOENG. 82:224-232(1996).  
 DR EMBL; D83793; D1021976; -.  
 KW SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 1 26 POTENTIAL.  
 SQ SEQUENCE 461 AA; 52213 MW; DF6927EF CRC32;

Query Match 9.5%; Score 368; DB 9; Length 461;  
 Best Local Similarity 34.88; Pred. No. 3.77e-49;  
 Matches 49; Conservative 41; Mismatches 49; Indels 2; Gaps 2;  
 Db 2 KKFVALLIMTFVVSMAVAQAASAKYSELEEGGVIMQAFYWDVPAAGIWDITIRSKIP 61  
 QY 5 NRIISVLLTLLAVAVLPYTPAQAHN-GTNGTMMQYFEWHL.PNDGNHNRRLRDA 63  
 Db 62 EWEYEGISAIWIPPAKGGGAYSMGYDYPDFDLGEYKQGTETFRFGSKELVRLIQ 121  
 QY 64 NLKSGITAVWIPPAWKGTSON-DVGYGAYDLYDLGEFNQKGTVTKYGTROLOQAVTS 122  
 Db 122 AHAYGKIVADIYVNHRRAGD 142  
 QY 123 LKNGIQVIGDVVNMHKGAD 143

RESULT 8  
 ID O08452 PRELIMINARY; PRT; 460 AA.  
 AC O08452;  
 DT 01-JUL-1997 (TREMBLREL, 04, CREATED)  
 DT 01-JUL-1997 (TREMBLREL, 04, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL, 05, LAST ANNOTATION UPDATE)  
 DE ALPHA AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN  
 GN AMYA.  
 OS PYROCOCUS FURIOSUS.  
 OC ARCHAEABACTERIA; CRENARCHAEOTA; THERMOPROTEALES; THERMOCOCCACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 3638;  
 RA DONG G., VIELLE C., SAVCHENKO A., ZEIKUS J.G.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM3638;  
 RX MEDLINE; 97341170.  
 RA JORGENSEN S., VORGAS C.E., ANTRANIKIAN G.;  
 RL J. BIOL. CHEM. 272:16335-16342(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM3638;  
 RA JOERGENSEN S.T., VORGAS C.E., ANTRANIKIAN G.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 DR EMBL; AF001268; G2183106; -.  
 DR EMBL; U96622; G2343083; -.  
 KW HYDROLASE; GLYCOSIDASE.  
 SQ SEQUENCE 460 AA; 52909 MW; EDEC186F CRC32;

Query Match 9.3%; Score 362; DB 9; Length 460;  
 Best Local Similarity 37.3%; Pred. No. 6.41e-48;  
 Matches 53; Conservative 36; Mismatches 51; Indels 2; Gaps 2;  
 Db 1 MNIKPLTLLTLLFFVLASPVSAKYALEEGGVIMQAFYWDVPGGIWHDHRSKI 59  
 QY 3 LHNRIISVLLTLLAVAVLPYTPAQAHN-GTNGTMMQYFEWHL.PNDGNHNRRLRDA 62  
 Db 60 PEWEYEGISAIWIPPAKGGGAYSMGYDYPDFDLGEYKQGTETFRFGSKELVRLIQ 119  
 QY 63 ANLKSIGITAVWIPPAWKGTSON-DVGYGAYDLYDLGEFNQKGTVTKYGTROLOQAVT 121  
 Db 120 TAHAYGKIVADIYVNHRRAGD 141  
 QY 122 LKNGIQVIGDVVNMHKGAD 143

RESULT 9  
 ID O60224 PRELIMINARY; PRT; 504 AA.  
 AC O60224;  
 DT 01-NOV-1996 (TREMBLREL, 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)  
 DE ALPHA-AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN  
 OS NATRONOCOCUS SP.  
 OC ARCHAEABACTERIA; EURYARCHAEOTA; HALOBACTERIALES; HALOBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AH-36;  
 RX MEDLINE; 94327485.  
 RA KOBAYASHI T., KANAI H., AONO R., HORIKOSHI K., KUDO T.;  
 RL J. BACTERIOL. 176:5131-5134(1994).  
 DR EMBL; D26510; G517137; -.  
 KW SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 1 43 POTENTIAL.  
 SQ SEQUENCE 504 AA; 55696 MW; A5E843BA CRC32;

Query Match 4.4%; Score 172; DB 9; Length 504;

[illegible]

SQ SEQUENCE 427 AA; 47456 MW; 2E2F49DA CRC32;

Query Match 3.5%; Score 135; DB 8; Length 427;  
Best Local Similarity 29.7%; Pred. No. 2.27e-05;  
Matches 30; Conservative 27; Mismatches 33; Indels 11; Gaps 6;

Db 27 LFOGFNWSKNGWTFNLMGKVDVDDTAAGVTHVWLPASQSVAE-Q-GYMPGRLYDL- 83  
QY 40 MMQYFEWHLPNDSGNHW-NRLRDDAANKSKGITAVWIPPAWKGTSDNDVGYGAYDLYDLG 98  
Db 84 D-----A--SKYGNKAOLKSLGALHGKGVKAIADIIVNHR 117  
QY 99 EFNQKGVTRTKYGRSLOQGAVTSLKNGIQVYGDVVVNNHK 139

RESULT 14  
ID Q41441 PRELIMINARY; PRT; 349 AA.  
AC Q41441;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE (FRAGMENT).  
GN AMY23.  
OS SOLANUM TUBEROSUM (POTATO).  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOITYLEDONEAE;  
OC SOLANALES; SOLANACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DIANELLA; TISSUE=SPROUT;  
RA GAUSING K., KREIBERG T.D.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; M81682; G1220451;  
DR EMBL; A21345; G579229;  
FT NON\_TER 1  
SQ SEQUENCE 349 AA; 39294 MW; 64DC2187 CRC32;

Query Match 3.4%; Score 130; DB 8; Length 349;  
Best Local Similarity 41.2%; Pred. No. 1.31e-04;  
Matches 21; Conservative 11; Mismatches 17; Indels 2; Gaps 2;

Db 223 PKRAVTFIDNHDGTGTONMPPPSDKVMQ-GYAYILTH-PGIPSVFYDHF 271  
QY 353 PIHAVTFVDNHDSPGGALESFVQSWFKPLAYALILTREQGYPSVYGDY 403

RESULT 15  
ID Q42504 PRELIMINARY; PRT; 421 AA.  
AC Q42504;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE TYPE A.  
GN AMY32B  
OS HORDEUM VULGARE (BARLEY).  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;  
OC CYPERALES; GRAMINEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WHITTIER R.F., DEAN D.A., ROGERS J.C.;  
RL NUCLEIC ACIDS RES. 13:0-0(1987).  
RN [2]  
RP SEQUENCE OF 1-46 FROM N.A.  
RX MEDLINE; 85006965.  
RA ROGERS J.C., MILLIMAN C.;  
RL J. BIOL. CHEM. 259:12234-12240(1984).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA WHITTIER R.F., DEAN D.A., ROGERS J.C.;  
RL SUBMITTED (FEB-1991) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; X05166; G19049;  
DR EMBL; M15208; G167001;  
SQ SEQUENCE 421 AA; 46102 MW; EAB3DF68 CRC32;

Query Match 3.4%; Score 132; DB 8; Length 421;  
Best Local Similarity 28.7%; Pred. No. 6.53e-05;  
Matches 29; Conservative 30; Mismatches 31; Indels 11; Gaps 6;  
Db 27 LFOGFNWSKNGWTFNLMGKVDVDDIAAGVTHVWLPASQSVAE-Q-GYMPGRLYDI- 83  
QY 40 MMQYFEWHLPNDSGNHW-NRLRDDAANKSKGITAVWIPPAWKGTSDNDVGYGAYDLYDLG 98  
Db 84 D-----A--SKYGNAAELKSLGALHGKGVKAIADIIVNHR 117  
QY 99 EFNQKGVTRTKYGRSLOQGAVTSLKNGIQVYGDVVVNNHK 139

Search completed: Tue Sep 15 14:21:46 1998  
Job time : 131 secs.



\*\*\*\*\*

W P S R L H  
 (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:16:37 1998; MasPar time 26.73 Seconds  
 Tabular output not generated. 705.206 Million cell updates/sec

Title: >US-08-952-741-2  
 Description: (1-516) from US08952741.pep  
 Perfect Score: 3873  
 Sequence: 1 MKLHNR-IISVLTLLAVAV.....ADGWGNTVNGGAVYVWKQ 516

Scoring table: PAM 150  
 Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: pir56

Statistics: 1.pir1 2.pir2 3.pir3 4.pir4 5.nrl3d

Mean 48.991; Variance 108.981; scale 0.450

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3368	87.0	518	1 A27705	alpha-amylase (EC 3.2	0.00e+00
2	2679	69.2	512	1 ALBSL	alpha-amylase (EC 3.2	0.00e+00
3	2661	68.7	549	1 A54541	alpha-amylase (EC 3.2	0.00e+00
4	2645	68.3	549	1 A24436	alpha-amylase (EC 3.2	0.00e+00
5	2623	67.7	549	1 A24549	alpha-amylase (EC 3.2	0.00e+00
6	2559	66.1	514	1 ALBSN	alpha-amylase (EC 3.2	0.00e+00
7	2524	65.2	548	1 ALBSF	alpha-amylase (EC 3.2	0.00e+00
8	1912	49.4	493	2 S15713	alpha-amylase (EC 3.2	0.00e+00
9	1558	40.2	290	5 1BPLB	alpha-1,4-glucan-4-gl	1.00e-259
10	1558	40.2	290	5 1VJS2	alpha-amylase (EC 3.2	1.00e-259
11	1403	36.2	494	1 B45738	alpha-amylase (EC 3.2	2.02e-230
12	1360	35.1	495	1 A45738	alpha-amylase (EC 3.2	2.62e-222
13	1006	26.0	179	5 1BPLA	alpha-1,4-glucan-4-gl	6.06e-156
14	1006	26.0	179	5 1VJS1	alpha-amylase (EC 3.2	6.06e-156
15	344	8.9	217	2 A19506	alpha-amylase (EC 3.2	1.87e-36
16	217	5.6	547	2 A32803	glucan 1,4-alpha-malt	7.94e-16
17	218	5.6	551	2 S05667	glucan 1,4-alpha-malt	5.58e-16
18	203	5.2	417	5 1AMG	1,4-alpha-d-glucan ma	1.08e-13
19	203	5.2	418	5 1JDD	1,4-alpha maltotetra	1.08e-13
20	203	5.2	418	5 1JDC	1,4-alpha maltotetra	1.08e-13
21	203	5.2	418	5 1JDA	1,4-alpha maltotetra	1.08e-13
22	203	5.2	528	1 ALBSK	alpha-amylase (EC 3.2	1.52e-13
23	192	5.0	713	2 S09196	alpha-amylase (EC 3.2	4.77e-12

24 189 4.9 712 1 ALBSG3 cyclomaltodextrin glu 1.33e-11  
 25 189 4.9 718 1 ALBSG6 cyclomaltodextrin glu 1.33e-11  
 26 186 4.8 684 5 1CGT cyclomaltodextrin glu 3.67e-11  
 27 186 4.8 686 5 1CDG cyclomaltodextrin glu 3.67e-11  
 28 185 4.8 686 5 1CDG cyclomaltodextrin glu 5.14e-11  
 29 185 4.8 686 5 1CXH cyclodextrin glycosyl 5.14e-11  
 30 185 4.8 686 5 1CXE cyclodextrin glycosyl 5.14e-11  
 31 185 4.8 686 5 1CXI cyclodextrin glycosyl 5.14e-11  
 32 185 4.8 686 5 1CXG cyclomaltodextrin glu 5.14e-11  
 33 185 4.8 686 2 S43819 cyclomaltodextrin glu 5.14e-11  
 34 185 4.8 686 5 1TCNA cyclodextrin glycosyl 5.14e-11  
 35 185 4.8 686 5 1TCMB cyclodextrin glycosyl 5.14e-11  
 36 186 4.8 718 1 ALBSG3 cyclomaltodextrin glu 3.67e-11  
 37 182 4.7 680 5 1CYG cyclodextrin glucanot 1.41e-10  
 38 183 4.7 686 5 1PAMA cyclodextrin glucanot 1.01e-10  
 39 183 4.7 686 5 1PAMB cyclodextrin glucanot 1.01e-10  
 40 181 4.7 686 5 1DIJ cyclodextrin glycosyl 1.98e-10  
 41 181 4.7 686 5 1CGV cyclomaltodextrin glu 1.98e-10  
 42 182 4.7 711 1 ALBSAF cyclomaltodextrin glu 1.41e-10  
 43 183 4.7 713 1 ALBSG7 cyclomaltodextrin glu 1.01e-10  
 44 183 4.7 713 1 ALBSG1 cyclomaltodextrin glu 1.01e-10  
 45 181 4.7 718 1 ALBSMX cyclomaltodextrin glu 1.98e-10

#### ALIGNMENTS

RESULT 1

ENTRY A27705 #type complete  
 TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus sp.  
 ALTERNATE\_NAMES 1,4-alpha-D-glucan glucanohydrolase; G6-amylase  
 ORGANISM #formal\_name Bacillus sp.  
 DATE 31-Mar-1989 #sequence\_revision 18-Aug-1995 #text\_change 05-Sep-1997  
 ACCESSIONS A27705  
 REFERENCE A27705  
 #authors Tsukamoto, A.; Kimura, K.; Ishii, Y.; Takano, T.; Yamane, K.  
 #journal Biochem. Biophys. Res. Commun. (1988) 151:25-31  
 #title Nucleotide sequence of the maltohexose-producing amylase gene from an alkalophilic Bacillus sp. 707 and structural similarity to liquefying type alpha-amylases.  
 #cross-references MUID:88162814  
 #accession A27705  
 #molecule\_type DNA  
 ##residues 1-518 ##label TSU  
 ##cross-references GB:M1862; NID:g142496; PID:g142497  
 ##experimental\_source chromosomal DNA of strain 707  
 ##note amino end of mature protein also determined  
 ##comment This is the smallest of five starch-hydrolyzing enzymes from this organism.

#### FUNCTION

#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 #pathway glycogen/starch degradation  
 #CLASSIFICATION #superfamily alpha-amylase, amyloliquefaciens type;  
 #pathway alpha-amylase core homology  
 #KEYWORDS extracellular protein; glycosidase; hydrolase; polysaccharide degradation  
 #FEATURES  
 1-33 #domain signal sequence #status predicted #label SIG\  
 34-518 #product alpha-amylase #status experimental #label MAT\  
 236-369 #domain alpha-amylase core homology #label AMY\  
 139,238,273 #binding\_site calcium (Asn, Asp, His) #status predicted  
 269,299,366 #active\_site Asp, Glu, Asp #status predicted  
 SUMMARY #length 518 #molecular-weight 59008 #checksum 7204

Query Match 87.0%; Score 3368; DB 1; Length 518;  
 Best Local Similarity 82.8%; Pred. No. 0.00e+00;  
 Matches 427; Conservative 57; Mismatches 30; Indels 2; Gaps 2;

Db 3 MRTGKGFISILLAFLLVTSI--PFTLVDEAHNGTNGTMMQYFEWYLPNDGHNHRLN 61

QY 1 MKLHNR-IISVLTLLAVAVLPYMTPEQAHHNGTNGTMMQYFEWYLPNDGHNHRLN 59

```

Db 62 SDASNLKSGKITAVIPPAWKGASQNDVGVGAYDLXDLGBFNOKGTVRTYKIGTRSQLOAA 121
QY 60 DDAANLKSAGITAVIPPAWKGTSQNDVGVGAYDLXDLGBFNOKGTVRTYKIGTRSQLOQA 119
Db 122 VTSLKNGIQVYGDVYVNNHKGADATEMVAVENPNRNNOETGTYTIEAWTRFDPGR 181
QY 120 VTSLKNGIQVYGDVYVNNHKGADGTEMVAVENPNRNNOESGTYTIEAWTRFDPGR 179
Db 182 GNTHSFKRWYHFDGVWDQSRRLNRIYKFRGHGKAWDEVDTEGNTDYLMYADIDM 241
QY 180 GNTHSFKRWYHFDGTDWDQSRQLNKIYKFRGTGKAWDEVDIENGNDYLMYADIDM 239
Db 242 DRPEVYNELRNWGWYTNLGLGDFRIDAVKHIIKYSTROWINHYRSATGKNMFAVAFW 301
QY 240 DRPEVINELRNWGWYTNLNLGDFRIDAVKHIIKYSTROWLTHVRNTTKGPMFAVAFW 299
Db 302 KNDLGAENYLOKTNHNSVEDYPLHNLNASKSGNDYMRINFNGTVVQRPSPHATVF 361
QY 300 KNDLAAIENLNTSNHNSVDFPLHNLNASKSGYFDMRILNLSGVVQKHPHATVF 359
Db 362 VDNHDSQPEALESFVEEFKPLAYALTITREQQPSVFYGYGIPTHGVPAMRSKIDP 421
QY 360 VDNHDSQPEALESFVQSNFKPLAYALLTREQQPSVFYGYGIPTHGVPMSKSKIDP 419
Db 422 ILEARKIYAGKQNDYLDHNNIIGTWREGTAHPNSGLATIMSDGAGGSKWFMVGRNKAG 481
QY 420 LLQARQTYAIGTDHDFDHDHIIIGTWREGDSHPNSGLATIMSDGPGGNKMWYVGRHKAG 479
Db 482 QVMSDITGNRTGVTINADGWNFSYNGSGSVIWN 517
QY 480 QVWRDITGNRSGVTINADGWNFTVNGGAVSVWK 515

RESULT 2
ENTRY ALBSL #type complete
TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus licheniformis
ALTERNATE_NAMES 1,4-alpha-D-glucan glucanohydrolase
ORGANISM #formal_name Bacillus licheniformis
DATE 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 20-Mar-1998
ACCESSION A91997; B24549; A91796; A21663; A26151; S53788; I39772;
I39774; A00844
REFERENCE A91997
#authors Yuuki, T.; Nomura, T.; Tezuka, H.; Tsuboi, A.; Yamagata, H.;
Tsukagoshi, N.; Uda, S.
#journal J. Biochem. (1985) 98:1147-1156
#title Complete nucleotide sequence of a gene coding for heat- and
pH-stable alpha-amylase of Bacillus licheniformis:
Comparison of the amino acid sequences of three bacterial
liquefying alpha-amylases deduced from the DNA sequences.
#cross-references MUID:86111694
#accession A91997
#molecule_type DNA
#residues 1-512 #label YUU
#cross-references GB:X03236; NID:g39551; PID:g39552
#experimental_source ATCC 27811
REFERENCE A91817
#authors Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindie,
K.L.; Carmona, C.; Requa, C.
#journal J. Bacteriol. (1986) 166:635-643
#title Structural genes encoding the thermophilic alpha-amylases of
Bacillus stearothermophilus and Bacillus licheniformis.
#cross-references MUID:86195857
#accession B24549
#molecule_type DNA
#residues 1-162, 'L', 164-338, 'G', 340-348, 'S', 350-512 #label GRA
#cross-references GB:M13256; NID:g142510; PID:g142511
#experimental_source NCIB 8061
REFERENCE A91796
#authors Stephens, M.A.; Ortlepp, S.A.; Ollington, J.F.; McConnell,
D.J.
#journal J. Bacteriol. (1984) 158:369-372
#title Nucleotide sequence of the 5' region of the Bacillus

```

```

licheniformis alpha-amylase gene: comparison with the
Bacillus amyloliquefaciens gene.
#cross-references MUID:84185455
#accession A91796
#molecule_type DNA
#residues 1-104 #label STE
#cross-references GB:K01984; NID:g142432; PID:g142433
REFERENCE A21663
#authors Sibakov, M.; Palva, I.
#journal Eur. J. Biochem. (1984) 145:567-572
#title Isolation and the 5'-end nucleotide sequence of Bacillus
licheniformis alpha-amylase gene.
#cross-references MUID:85076654
#accession A21663
#molecule_type DNA
#residues 1-3, 'H', 5-12, 'P', 14-47, 'R', 49-61, 'V', 63, 'D', 65-67, 'VA',
70-71, 'S', 73-80, 'D', 82-104, 118-121 #label STB
#experimental_source chromosomal DNA of ATCC 14580
#note the authors translated the codon CGT for residue 48 as
Gly and GAC for residue 64 as His
REFERENCE A26151
#authors Kuhn, H.; Fietzek, P.; Lampen, J.O.
#journal J. Bacteriol. (1982) 149:372-373
#title N-terminal amino acid sequence of Bacillus licheniformis
alpha-amylase: comparison with Bacillus amyloliquefaciens
and Bacillus subtilis enzymes.
#cross-references MUID:82098050
#accession A26151
#molecule_type protein
#residues 30-37, 'E', 39-41, 'X', 43-47 #label KUH
REFERENCE S53788
#authors Machius, M.; Wiegand, G.; Huber, R.
#journal J. Mol. Biol. (1995) 246:545-559
#title Crystal structure of calcium-depleted Bacillus licheniformis
alpha-amylase at 2.2 A resolution.
#accession S53788
#molecule_type protein
#residues 'D', 220-227 #label MAC
#note sequence represents amino end of an internal fragment
created by a single enzymatic cleavage by a protease
trace contaminant during purification
REFERENCE I39772
#authors Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.
#journal Gene (1990) 96:37-41
#title In vivo genetic engineering: homologous recombination as a
tool for plasmid construction.
#cross-references MUID:91092499
#accession I39772
#status translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 1-32, 'I' #label RES
#cross-references GB:M62637; NID:g142498; PID:g142499
REFERENCE I39773
#authors Laoide, B.M.; Chambliss, G.H.; McConnell, D.J.
#journal J. Bacteriol. (1989) 171:2435-2442
#title Bacillus licheniformis alpha-amylase gene, amyL, is subject
to promoter-independent catabolite repression in Bacillus
subtilis
#cross-references MUID:89213924
#accession I39774
#status translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 1-29 #label RE2
#cross-references GB:M26412; NID:g341477; PID:g516590
GENETICS amyL
#gene catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
#CLASSIFICATION #superfamily alpha-amylase, amyloliquefaciens type;
alpha-amylase core homology
#KEYWORDS extracellular protein; glycosidase; heat-stable protein;

```

FUNCTION	#start_codon	GTG	#description						
CLASSIFICATION	#pathway	glycogen/starch degradation	catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds						
KEYWORDS	#superfamily	alpha-amylase, amyloliquefaciens type;							
FEATURE	#alpha-amylase core homology	extracellular protein; glycosidase; heat-stable protein; hydrolase; polysaccharide degradation							
1-34	#domain signal sequence	#status predicted	#label SIG\						
35-549	#product	alpha-amylase	#status predicted	#label MAT\					
235-368	#domain	alpha-amylase core homology	#label AMY\						
139,237,272	#binding_site	calcium (Asp, Asp, His)	#status predicted\						
268,298,365	#active_site	Asp, Glu, Asp	#status predicted						
SUMMARY	#length	549	#molecular-weight	62598	#checksum	5758			
Query Match	68.7%	Score	2661;	DB	1;	Length	549;		
Best Local Similarity	67.5%	Pred. No.	0.00e+00;						
Matches	341;	Conservative	78;	Mismatches	81;	Indels	5;	Gaps	5;
Db	15	LLAFLLTVS-LFCPTGQPAKAA-APFNCTMMQYFWYLDDCTGLTWKVANEANNSSLGI	72						
Qy	11	LLTLLAVAVLPYPTFAQAHHNCTNTMMQYFWHLPLDNCNHNRLRDDANLAKSKI	70						
Db	73	TALMLPPAYKTRSDVGYGYDLYDLGEFNOGKGVTRTKYGTAKQYLO-AIQAAHAAGMQ	131						
Qy	71	TAVMIIPAAWKTQNDVGYGAYDLYDLGEFNOGKGVTRTKYGTQSQ-LOGAVTSLKNGIQ	129						
Db	132	VYADVDFHKGADGTETWDAVEVNPDRNQIEISGTQYIQAWTKDFPGRGNTSYSFKWR	191						
Qy	130	VYGDVYMHKGGADGTEMVNAVEVNRNRNQIEISGTYTTEAWTKFDFPGRGNTHSNFKWR	189						
Db	192	WYHFDGVDWDSRKL-SRIYKFGTCKAWDWEVDFTENGNDYLYMADLMDHDPVYTELK	250						
Qy	190	WYHFDGTWDQSRQLNKIYKFGTCKAWDWEVDLENGNDYLYMADLMDHDPVYTELK	249						
Db	251	NWGKVVYVNTTIDGRDLDAVAKHIKEFFPDWLSYVRSOTGKPLFTVGEYSWSDINKLHY	310						
Qy	250	NWGYYVNTNLDFRIDAVAKHIKSYTFDWTHTVNTTGTGKPMFAVEFKNDLAAIENY	309						
Db	311	ITKTDGTVMSLFDAPLHNKFFYTASKSGGFADMTLMTNTLMKDQPTLAVTFVDNHDTEPGQ	370						
Qy	310	LNKTSWNHSVDFVPLHYNLNASNGGYFDMRNILNGSVVQKHPHAVTVDNHDSPQGE	369						
Db	371	ALQSVWDFWFXPLAFILTRQEGYPCFYGYGIPQYINISLKSIDPLLIARDIYAY	430						
Qy	370	ALESEVQSWFXPLAYALILTRQEGYPCFYGYGIPTHGVPSMKSIDPLLIQAROTIYAY	429						
Db	431	GTQHDYLDHSDIIGWTRGGTEKPCGSLAALITDGPGGSKMVMYVGKHAGKVFYDUTGNR	490						
Qy	430	GTQHDYFDHHDILGWTRGDSHPNSGLATIMSDGPGGNKVMYVGKHAGKQVWRDITGNR	489						
Db	491	SDTVTINSDGWGEFKVNGSGSVVW	515						
Qy	490	SGTVTINADGNGNFTVNGAVSVVW	514						
RESULT	4								
ENTRY	A24436	#type	complete						
TITLE	alpha-amylase (EC 3.2.1.1)	precursor	- Bacillus						
ALTERNATE_NAMES	stearothermophilus plasmid pat5								
ORGANISM	1,4-alpha-D-glucan glucanohydrolase								
DATE	#formal_name	Bacillus stearothermophilus							
	05-Jun-1987	#sequence_revision	18-Aug-1995	#text_change					
	29-Aug-1997								
ACCESSIONS	A24436;	I39777							
REFERENCE	A24436								
#authors	Nakajima, R.; Imanaka, T.; Aiba, S.								
#journal	J. Bacteriol. (1985)	163:401-406							
#cross-references	MUID:85234394								
#accession	A24436								
#molecule	type	DNA							

```
##residues 1-549 #label NAK
##cross-references GB:M1450
##experimental_source plasmid pAT5
##note
amino end of the mature protein also determined
REFERENCE
I39772
Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.
Gene (1990) 96:37-41
In vivo genetic engineering: homologous recombination as a
tool for plasmid construction.
#cross-references MUID:91092499
#accession I39777
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
##residues 1-45 #label RES
##cross-references GB:M62638; NID:g142514; PID:g142515
COMMENT
Alpha-amylase genes have been found on plasmids and in multiple
copies on the chromosome in various strains of this organism.
GENETICS
#gene amyS
#genome plasmid
#start_codon GTG
FUNCTION
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
CLASSIFICATION
#superfamily alpha-amylase, amylioliquefaciens type;
alpha-amylase core homology
extracellular protein; glycosidase; heat-stable protein;
hydrolase; polysaccharide degradation
FEATURE
1-34 #domain signal sequence #status predicted #label SIG\
35-549 #product alpha-amylase #status experimental #label MAT\
235-368 #domain alpha-amylase core homology #label AMY\
139,237,272 #binding_site calcium (Asp, Asp, His) #status predicted\
268,298,365 #active_site Asp, Glu, Asp #status predicted
SUMMARY
#length 549 #molecular-weight 62670 #checksum 5048
Query Match 68.3%; Score 2645; DB 1; Length 549;
Best Local Similarity 67.1%; Pred. No. 0.00e+00;
Matches 339; Conservative 81; Mismatches 80; Indels 5; Gaps 5;
Db 15 LIAFLLT-ALLFPCPTGCPAKAA-APFNGTMMQYFEWYLPDDGTLTKVANEANNLSSLGI 72
Qy 11 LTLTLLAVAVLPFYTPTEPAQAHHNGTNGTMMQYFEWHLPNDRHNNRLRDAANLKSXGI 70
Db 73 TALWLPAYKGRSDVGVVDLGLDFNFKGAVRKYGTAKYQLQ-AIQAAHAAGMQ 131
Qy 71 TAVWIPPAWKGTSDNDVGYGAYDLGLDFNFKGAVRKYGTAKYQLQ-LOGAVTSLKXNGIQ 129
Db 132 VYADVDFHKGAGDTEWYDAVEVNPDRNQISGTGYQIAWTKDFPGRGNTYSSFKWR 191
Qy 130 VTGDMVNMHKGADGTWYNAVEVNRNRNQISGEYTIKATKDFPGRGNTSHNFKWR 189
Db 192 WYHFDGVDWDESKL-SRIYKFRGIGKAWDWEVDENGNYDILMYADLMDHPEVVTEUK 250
Qy 190 WYHFDGTDWQSKLQNKIYKFRGTGKAWDWEVDENGNYDILMYADLMDHPEVINELR 249
Db 251 SNGKVVNTNIDGFRDLDAVKIKFSFFPDWLSVRSOTGKPLFTVGYEWSYDINKLHNY 310
Qy 250 NMGVWYNTNLNDGFRIDAVKHKYSYTRDNLTHVRNTTGKMFVFAEFAEKKNDLAAIENY 309
Db 311 IMKTNGTMSLDAPLHNKYTASKSGTGTDMRTLTMTNLTKMDQPTLAVTFVDNHDPEPGQ 370
Qy 310 LNKTSNNHSDVDVPLHYNLYNASNSGDFYDMRNLNGSVVQKHPIHAVTFVDNHDSPQGE 369
Db 371 ALOSVDWPFKPLAYAFILTRQEGYPCVYGYDYGIPQINIPSLKSIDPLLIARDYAY 430
Qy 370 ALESFQSVFKPLAYALILTRQEGYPCVYGYDYGIPQINIPSLKSIDPLLIARDYAY 429
Db 431 GTQHDYLDHSDIIGWTRREGVTEKPGSGLAALITDGGGSKWYVGVQKHAGKVFDITGNR 490
Qy 430 GTQHDYLDHSDIIGWTRREGVTEKPGSGLAALITDGGGSKWYVGVQKHAGKVFDITGNR 499
```

```
Db 491 SDVTINSDGWGEKFGVNGSVSVWV 515
Qy 490 SGVTIINADGWNFTVNGGAVSVWV 514

RESULT 5
ENTRY A24549 #type complete
TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus
steartothermophilus (strain NZ-3)
ALTERNATE_NAMES 1,4-alpha-D-glucan glucanohydrolase
ORGANISM #formal_name Bacillus steartothermophilus
DATE 30-Jun-1988 #sequence_revision 18-Aug-1995 #text_change
05-Sep-1997
ACCESSIONS A24549; I39501; I39770
REFERENCE A91817
#authors Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindle,
J. Bacteriol. (1986) 166:635-643
#journal Structural genes encoding the thermophilic alpha-amylases of
#title Bacillus steartothermophilus and Bacillus licheniformis.
#cross-references MUID:86195857
#accession A24549
#molecule_type DNA
#residues 1-549 #label GRA
##cross-references GB:M13255; NID:g142512; PID:g142513
##experimental_source genomic DNA of strain NZ-3
REFERENCE I39501
#authors Satoh, H.; Nishida, H.; Isono, K.
#journal J. Bacteriol. (1988) 170:1034-1040
#title Evidence for movement of the alpha-amylase gene into two
phylogenetically distant Bacillus steartothermophilus
strains.
#cross-references MUID:88139156
#accession I39501
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 536-549 #label RES
##cross-references GB:M29577; NID:g142476; PID:g142478
##experimental_source strain DY-5
#accession I39770
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 536-549 #label RE2
##cross-references GB:M29578; NID:g142484; PID:g142486
##experimental_source strain 799
COMMENT Alpha-amylase genes have been found on plasmids and in multiple
copies on the chromosome in various strains of this organism.
GENETICS
#start_codon GTG
FUNCTION
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
CLASSIFICATION
#superfamily alpha-amylase, amylioliquefaciens type;
alpha-amylase core homology
extracellular protein; glycosidase; heat-stable protein;
hydrolase; polysaccharide degradation
KEYWORDS
FEATURE
1-34 #domain signal sequence #status predicted #label SIG\
35-549 #product alpha-amylase #status predicted #label MAT\
235-368 #domain alpha-amylase core homology #label AMY\
139,237,272 #binding_site calcium (Asp, Asp, His) #status predicted\
268,298,365 #active_site Asp, Glu, Asp #status predicted
SUMMARY
#length 549 #molecular-weight 62643 #checksum 6769
Query Match 67.7%; Score 2623; DB 1; Length 549;
Best Local Similarity 66.9%; Pred. No. 0.00e+00;
Matches 338; Conservative 79; Mismatches 83; Indels 5; Gaps 5;
Db 15 LIAFWLTAS-LFCPTGCPAKAA-APFNGTMMQYFEWYLPDDGTLTKVANEANNLSSLGI 72
Qy 11 LTLTLLAVAVLPFYTPTEPAQAHHNGTNGTMMQYFEWHLPNDRHNNRLRDAANLKSXGI 70
```



Db	131	QVTVADVYDFDHKGGADGTWVDAVENPNSDRNOEISGTYIOAWTKFDPGRCNTYSSFKW	190
Qy	129	QVYGVVNMHKGKGGADGTWVDAVENPNSDRNOEISGTYIOAWTKFDPGRCNTYSSFKW	188
Db	191	RWYHFDGVDWDESRL-SRIYKFRGIGKAWDEVDTENGYDYLMAVDLMDHPEVVTIEL	249
Qy	189	RWYHFDGTDWQSOLOQNKIYKFGTGKAWDEVDTENGYDYLMAVDLMDHPEVINEL	248
Db	250	KNMGKWTVNTNIDGFRDLGLKHIFSEFPDWLSVRSQTKPLFTVGVYSYDINKLHN	309
Qy	249	RNMGWITNTLNLGDFRIDAVKHIKSYTRDLWTLVTRNTTGKPMFAVAFWKNDLAIEN	308
Db	310	YITKNTGMSLFDAPLHNKFTASKSGAFDNRITLMTNLMDKDOPTLAVTFVDNHDINPA	369
Qy	309	YLNKTSWNHVSFVPLHYLNTNLSNGGVDNRNLNLSGVQKHPTLAVTFVDNHDISQPG	368
Db	370	KRC-SHGPRWPKPLAYATILRQEGYPCVFGYDYGIPQYNIPSLKSKIDPLLIARRDYA	428
Qy	369	EALSEFVOSWFKPLAYALLRQEGYPSVFGYDYGIPGTHGVPSMKSKIDPLLOARQTYA	428
Db	429	YGTQHDYLDHSDIIGWTRREGVTEKPGSGUAAITDCAGRSKMYGVQKHAGKVFYDLTGN	488
Qy	429	YGTQHDYFDHDIIGWTRREGDSHNSGLATIMSDPGGKNMYGVQKHAGQVWVRDITGN	488
Db	489	RSDTVTINSNGWGEFKVNGSGVSVMV	514
Qy	489	RSQVTINADGNGTVNGGVSVMV	514
RESULT	8	#type complete	
ENTRY	S15713	#formal_name Bacillus circulans	
TITLE	S15713	#formal_name Bacillus circulans	
ORGANISM	S15713	#sequence_revision 10-Nov-1995	#text_change
DATE	08-Sep-1997		
ACCESSIONS	S15713		
REFERENCE	Marcel, T.		
#authors	submitted to the EMBL Data Library, May 1991		
#submission	S15713		
#accession	S15713		
#molecule_type	DNA		
GENETICS	1-493	#label MAR	
#residues	EMBL:X60779; NID:g39411; PID:g39412		
#cross-references			
#gene	amyE		
FUNCTION	catyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds		
#description	glycogen/starch degradation		
#pathway	#superfamily alpha-amylase, amyloliquefaciens type;		
CLASSIFICATION	alpha-amylase core homology		
KEYWORDS	glycosidase; hydrolase; polysaccharide degradation		
FEATURE	#domain alpha-amylase core homology #label AMY		
200-333	#length 493 #molecular-weight 56536 #checksum 8821		
SUMMARY			
Query Match	49.4%; Score 1912; DB 2; Length 493;		
Best Local Similarity	51.5%; Pred. No. 0.00e+00;		
Matches	245; Conservative 94; Mismatches 135; Indels 2; Gaps 2;		
Db	6	TMMOFFEWHLAAGDHWKRLAEMAPELKAGIDTVWVPVTKAVSAEDTGYGYDLYDLG	65
Qy	39	TMQIFEWHLPLNDGNHNNRLRDDAANKSGITAVWPPAKMGTSONDVGYGAYDLYDLG	98
Db	66	EFDQGTVRTYGTQBELBAIAECQKNGIAVYVDLVNHNKAGADETEVKFVIEVDPNDR	125
Qy	99	EFNQGTVRTYKGTQSLQGAVTSLKNNIGIQVGDVNMHKGKGGADGTETMNAVEVNRNR	158
Db	126	TKEISEPFELEGWTKFPGRCGOYSFSPKWNSEHFNQDDEF-AREERTGVFRIGAKKKW	184
Qy	159	NQELISGTYTEAWTKFDPGRCNTYSSFKWRYHFDGTDWQSOLOQNKIYKFGTGKAW	218

Db 185 NENVDFEGNXYLMEANIDYNNHPDVRREMDWGKWLIDTLQCGFRDLDAIKHINHEFIK 244  
QY 219 DWEVDIENGNYDLYMAYADIMDHPVEINELRNWGWYVNTLNLDGFRIDAYKHIKYSYTR 278  
Db 245 EFAAEMIRKRGQDFYIVGEFWSNLDACRFFLDVQIDFLDVSJLHYKLEASLKGGRDF 304  
QY 279 DWLTHVRNTTGKPMFAEWFKNLDAIENLYNKTNSHNSVDFVPLHYNLYNASNGGYF 338  
Db 305 DLSKIFDDTLVQTHPTHAVTFVNDHDSQPHEALESWIGDMFKPSAYALTLLRRDGYPVVF 364  
QY 339 DMENILNGSVVQKPIHAVTFVNDHDSQPGEALESFVQSFVKPLAYALILTRQGGYPSVF 398  
Db 365 YGDIYIGGPEPDGKKEIIDLARSARCWAYGEQEDYFDHANTIGVWRVVEIEGSC 424  
QY 399 YGDIYIGP-THGVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRREGDSSHPNSGL 457  
Db 425 AVVISNGDDGKRMFGEHRAGEVWDLTKSCDDQITIEEDGWATFHVCGGGSVW 480  
QY 458 ATMSDGPCKWYVYKHKAGQWRDITGNRSCTVTINADGNGFTVNGGAVSVW 513

RESULT 9  
ENTRY 1BPLB #type complete  
TITLE alpha-1,4-glucan-4-glucanohydrolase (EC 3.2.1.1), chain B -  
ALTERNATE\_NAMES Bacillus licheniformis  
PDB\_TITLE alpha-amylase (bla)  
ORGANISM glycosyltransferase  
#note #formal\_name Bacillus licheniformis  
REFERENCE ATCC: 27811  
#authors A65206  
#submission Machius, M.; Wiegand, G.; Huber, R.  
#cross-references PDB:1BPL submitted to the Brookhaven Protein Data Bank, July 1995  
REFERENCE S53788  
#authors Machius, M.; Wiegand, G.; Huber, R.  
#journal J. Mol. Biol. (1995) 246:545-559  
#title Crystal structure of calcium-depleted Bacillus licheniformis  
alpha-amylase at 2.2 A resolution.  
COMMENT Resolution: 2.2 angstroms  
COMMENT Determination: X-ray diffraction  
COMMENT R-value: no refinement  
KEYWORDS 4-glucan-4-glucanohydrolase; alpha-1; alpha-amylase  
glycosyltransferase; glycosyltransferase  
FEATURE  
2-5 #region helix (right hand 3-10)\\  
14-31 #region helix (right hand 3-10)\\  
41-43 #region helix (right hand 3-10)\\  
46-60 #region helix (right hand 3-10)\\  
75-84 #region helix (right hand 3-10)\\  
94-105 #region helix (right hand 3-10)\\  
112-116 #region helix (right hand 3-10)\\  
120-123 #region helix (right hand 3-10)\\  
125-127 #region helix (right hand 3-10)\\  
149-161 #region helix (right hand 3-10)\\  
171-175 #region helix (right hand 3-10)\\  
189-201 #region helix (right hand 3-10)\\  
250-252 #region helix (right hand 3-10)\\  
35-38,65-68 #region beta sheet\\  
207-210,216-221, #region beta sheet\\  
232-237,285-288 #region beta sheet\\  
242-247,275-280 #region beta sheet\\  
7-9 #region beta sheet\\  
39,69,136 #site Asp, Glu, Asp #label ACT  
SUMMARY #length 290 #molecular-weight 33024 #checksum 5042  
Query Match 40.2%; Score 1558; DB 5; Length 290;  
Best Local Similarity 67.5%; Pred. No. 1.00e-259;  
Matches 195; Conservative 51; Mismatches 40; Indels 3; Gaps 2;  
Db 1 YDLYMADIDYDHPDVAEIKRWGTWYANELQDGLDFRLDAVKHKIFSLDRWVNHVREKT 60  
QY 229 YDLYMADIDMDHPVEINELRNWGWYVNTLNLDGFRIDAYKHIKYSYTRDWLTHVRNTT 288

Db 61 GKEMFTVAEYWDNDLGALENYLNKTNFNHSVDFVPLHYQFHAASSTOGGDMRKLNSTV 120  
QY 289 GKPMFAEAFKNDLAAIENLYNKTNSHNSVDFVPLHYNLYNASNGGYDFMRNLGVS 348  
Db 121 VSKHPLKAVTFVNDHDTQPGQSLSTVQWFKPLAYAFILTRREGYGPVFGMYGTHKGD 180  
QY 349 VOKHPIHAVTFVNDHDSQPGEALESFVQSFVKPLAYALILTRQGGYPSVFYDYG 406  
Db 181 SOREIPALKHKTEPILKARKQYAGAOHDYFDHHDIVGWTREGDSSVANSGLAALITDGP 240  
QY 407 TH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRREGDSSHPNSGLATIMSDGP 465  
Db 241 GGAKRMVYGRQAGTHTDITNRRSEPVVINSEGWGEFHVNGGSVIYV 289  
QY 466 GGNKWMYVGGKHKAGQWRDITGNRSCTVTINADGNGFTVNGGAVSVW 514

RESULT 10  
ENTRY 1VJS2 #type fragment  
TITLE alpha-amylase (EC 3.2.1.1), fragment 2 - Bacillus  
ALTERNATE\_NAMES bla  
PDB\_TITLE structure of alpha-amylase precursor  
ORGANISM #formal\_name Bacillus licheniformis  
#note strain bacillus licheniformis, ATCC: 27811  
REFERENCE A66860  
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#submission submitted to the Brookhaven Protein Data Bank, October 1996  
#cross-references PDB:1VJS  
REFERENCE TN032246  
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#book In Enzymes for Carbohydrate Engineering (In: Prog. Biotechnol., V.12), Park, K.H.  
#title Crystal structure of bacillus licheniformis alpha-amylase at 1.7 A resolution.  
REFERENCE TN032247  
#authors Lee, S.Y.; Kim, S.; Sweet, R.M.; Suh, S.W.  
#journal Arch. Biochem. Biophys. (1991) 291:255  
#title Crystallization and a preliminary X-ray crystallographic study of alpha-amylase from bacillus licheniformis.  
COMMENT Resolution: 1.7 angstroms  
COMMENT Determination: X-ray diffraction  
COMMENT R-value: 0.199  
KEYWORDS carbohydrate metabolism; glycosidase; hydrolase  
FEATURE  
2-5 #region helix (right hand 3-10)\\  
14-31 #region helix (right hand 3-10)\\  
41-43 #region helix (right hand 3-10)\\  
46-60 #region helix (right hand 3-10)\\  
75-84 #region helix (right hand 3-10)\\  
94-105 #region helix (right hand 3-10)\\  
112-116 #region helix (right hand 3-10)\\  
120-123 #region helix (right hand 3-10)\\  
125-127 #region helix (right hand 3-10)\\  
149-161 #region helix (right hand 3-10)\\  
171-175 #region helix (right hand 3-10)\\  
189-201 #region helix (right hand 3-10)\\  
250-252 #region helix (right hand 3-10)\\  
129-131,166-170, #region beta sheet\\  
35-38,65-68 #region beta sheet\\  
207-210,216-221, #region beta sheet\\  
232-237,285-289 #region beta sheet\\  
242-247,275-280 #region beta sheet\\  
7-9 #region beta sheet\\  
SUMMARY #length 290 #checksum 5042  
Query Match 40.2%; Score 1558; DB 5; Length 290;  
Best Local Similarity 67.5%; Pred. No. 1.00e-259;  
Matches 195; Conservative 51; Mismatches 40; Indels 3; Gaps 2;  
Db 1 YDLYMADIDYDHPDVAEIKRWGTWYANELQDGLDFRLDAVKHKIFSLDRWVNHVREKT 60  
QY 229 YDLYMADIDMDHPVEINELRNWGWYVNTLNLDGFRIDAYKHIKYSYTRDWLTHVRNTT 288

QY	229	YDLYMTADIDMDHP	EVINELRNKGVWYNTNLNLDGFRIDAVKHKYSYTRDMLTHVRNTT	288
Db	61	GKEMFTVAYEYQND	LGALENLYNKTNFNSVFDVPLHVFQHAASSTOGGDMKRLNSV	120
QY	289	GKMPFAVAFWKND	LAAIENLYNKTSNHSVFDVPLHLYNASSNGSYFDMRNLNGSV	348
Db	121	YSKHLKAVTFVDN	HDHDTQPOSLESTVQTFWKFLAYAFILTRRESGYPOVPYGDMTGTGKD	180
QY	349	VOKHPHIAVTFVDN	HDHDSQGEALSFVSQWFKFLAYALILTRQGYPSFYGDYGI--P	406
Db	181	SQREIPALKHKEPIL	KARKQYAYGAQHDYFDHHDIVGWTRCEDSVANSGLAALITDGP	240
QY	407	TH-GVPSMSKSDP	LQARQYAYGTQHDYFDHHDILGWTRCEDSHSPNSGLATIMSDGP	465
Db	241	GGAKRWYGRQAGET	HDHDTGNRSEPVNLNSGEGEFHWGSGYSIYV	289
QY	466	GGNKWYGVCKHAGW	VRDITGNRSGVTIINADGNGFTVNGAVSVV	514
RESULT	11	B45738	#type complete	
ENTRY		alpha-amylase (EC 3.2.1.1),	cytoplasmic - Salmonella	
TITLE		typhimurium		
ALTERNATE_NAMES		1,4-alpha-D-glucan	glucanohydrolase	
ORGANISM		07_Apr-1994	#sequence_revision 18-Aug-1995	#text_change
DATE		05-Sep-1997		
ACCESSIONS		B45738		
REFERENCE		A45738		
AUTHORS		Raha, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, J. Bacteriol. (1992) 174:6644-6652		
JOURNAL		Escherichia coli produces a cytoplasmic alpha-amylase, amyA		
TITLE		B45738		
ACCESSION		#molecule_type DNA		
RESIDUES		#residues 1-494	#label RAH	
CROSS-REFERENCES		#cross-references GB:L01643; NID:g154043; PID:g154045		
GENETICS				
GENE				
FUNCTION				
DESCRIPTION				
PATHWAY				
CLASSIFICATION				
KEYWORDS				
FEATURE				
SUMMARY				
QUEST MATCH				
BEST LOCAL SIMILARITY				
MATCHES				
Db	3	NPTLQYFHWYYPDGK	LSELAERADGLNIDIGINMVLPPACKGASGYSYGYDTYDLF	62
QY	37	NGTMQYFEWHL	PNDGNHNRLLRDDAANKSKGITAWIPPAWKGSQN-DYGYGADLY	95
Db	63	DLGEFOCKGTATKY	GDKRQLITADLAKNNNTAVLLDVVYVNHKMGADKEKIRKRVNQ	122
QY	96	DLGEFNQGTATKY	GSQSQGAVTSILKNNGLQVTDVYVNHKMGADGTEWVNAEVR	155
Db	123	DRTQIDDDNIECE	GWTRYTFPAPAGYSNFIWD-YHCFSGIDHLENPD-EDGIFKIVND	180
QY	156	SRNNOEISGEY	TEIAWTKDFPFGNGTHSNFKRWYH-FDGTDMDDQSRQJONKIYKF-RG	213
Db	181	YTGQWNDVDN	DENGFDYILMGENIDFRNHAVTEIKYVARSWVWEQTHCDGFRDLAVRHI	240
QY	214	TGKANDWEVD	NIENGNDYILMTADIMDHPEVIELNRNGWVTNTNLNLDGFRIDAVKHI	272
Db	241	PAWYFKWIEHW	QAVAPKPLFIVAEYWSHEVDKLTQIVDQDKTMLFDAPLQMFHEAS	300



QY 37 NGTMMQYFEWHPNDGQHWKRLQNDAAANLSKSGITAVWIPPAWKGTSON-DVGSGAYDLY 95  
Db 63 DLGEFDQKSIPTKYGKDAQLAALDAIKRNDIAVLDDVVVNNHKGADKEAIRVQRVNA 122  
QY 96 DLGEFNQKGTVRTKYGTRSQLOGAVTSLKNGGQVGVVNNHKGADGTEMVNAVEVNR 155  
Db 123 DRTQ-IDEIIECEGWTRTFPARAQYQSFQIWDKFCFSIDHINPD-EDGIFKIVND 180  
QY 156 SNRQETSGEYTI-EAWTKEDFPGRGNTHSNFKRWYHFDGTDQDQSLQNLKIKYKF-RG 213  
Db 181 YTEGNDQVDDELGNFDYILNGENIDFRNHAUTEEIKYARWVMEQTOCDGFRLDVAKHI 240  
QY 214 -TKANDWEVDIENGNTDIYLAIDMDHPEVINELNRNGVWYTNLNLGDFRIDAVAKHI 272  
Db 241 PAWFYKEWIEHGVAPKPLFIYAETWSHEVDKLTQYIDQVEGKTMFLDAPLQMKFHEAS 300  
QY 273 KYSYTRDLWTHVRTTKPFAVAEFAEKNDLAAENLNKTSWNHSDVDFPLHNLINAS 332  
Db 301 RMGRDYDMTOIFGTLVEADPFHATVILVANDHOTOPLQALEAPVPEWPKPLAYAILIREN 360  
QY 333 NSGGYFDMRNLNGSVYQKPIHATFVNDHDSQGEALSFVQSWFKPLAYAILILREQ 392  
Db 361 GVPSEVPDLYGAHYEDVGGDQYIPDMPIIEQLDELILARQFAHGVQTLFDFHPNCI 420  
QY 393 GYPSVFYGD-Y---Y-GIPTHG-V-P-SMK--SKIDPLQARQTYAYGTOHDYFDHDI 442  
Db 421 AFRSGYDFEPGCG-VV-VMSGNDGEGKTHILGENYGNKTRDFLGNROERVVITDENGAT 478  
QY 443 GWTREGDSSHPNSGLATIMSDGPGKNWYVGNKAKQVWRDITGNRSGTGTINADGWN 502  
Db 479 FFCNGGSVSVWV 490  
QY 503 FTVNGGAVSVWV 514

RESULT 13  
ENTRY TITLE  
ALTERNATE\_NAMES 1BPLA #type complete  
PDB\_TITLE alpha-1,4-glucan-4-glucanohydrolase (EC 3.2.1.1), chain A -  
ORGANISM Bacillus licheniformis  
#note glycosyltransferase  
#authors #formal\_name Bacillus licheniformis  
#cross-references ATCC: 27811  
#submission Machius, M.; Wiegand, G.; Huber, R.  
#cross-references PDB:1BPL submitted to the Brookhaven Protein Data Bank, July 1995  
#authors S53788  
#journal  
#title  
COMMENT Resolution: 2.2 angstroms  
COMMENT Determination: X-ray diffraction  
KEYWORDS 4-glucan-4-glucanohydrolase; alpha-1; alpha-amylase  
glycosyltransferase; glycosyltransferase  
FEATURE  
19-25 #region helix (right hand alpha)\  
27-32 #region helix (right hand alpha)\  
78-91 #region helix (right hand alpha)\  
155-157 #region helix (right hand alpha)\  
5-7,37-39,94-99 #region beta sheet\  
131-135,111-117, #region beta sheet\  
173-175 #region beta sheet\  
158-161 #region beta sheet\  
SUMMARY #length 179 #molecular-weight 20547 #checksum 1012

Query Match 26.0%; Score 1006; DB 5; Length 179;  
Best Local Similarity 70.9%; Pred. No. 6.06e-156;  
Matches 127; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

Db 2 NGTLMQYFEWTPNDGQHWKRLQNDAAANLSKSGITAVWIPPAWKGTSONDVGAYDLYD 61  
QY 37 NGTMMQYFEWHPNDGQHWKRLQNDAAANLSKSGITAVWIPPAWKGTSONDVGAYDLYD 96  
Db 62 LGEPHQKGTVRTKYGKELQSAIKSLHSRDIINVYGDVVVNNHKGADGTEMVNAVEVDP 121  
QY 97 LGEPHQKGTVRTKYGKELQSAIKSLHSRDIINVYGDVVVNNHKGADGTEMVNAVEVNR 156  
Db 122 DNRNVISGEHLIKANTHFFPGRGSTYSDFKWHWHYHFDGTDWDESRLK-NRIYKFOGKA 179  
QY 157 NRNOEISGEYTIETAWTKEDFPGRGNTHSNFKRWYHFDGTDWDSRLQNLKIKYKFRGTG 215

RESULT 14  
ENTRY TITLE  
ALTERNATE\_NAMES 1VJS1 #type fragment  
PDB\_TITLE alpha-amylase (EC 3.2.1.1), fragment 1 - Bacillus  
ORGANISM licheniformis  
#note structure of alpha-amylase precursor  
#formal\_name Bacillus licheniformis  
#cross-references strain bacillus licheniformis, ATCC: 27811  
#authors A6860  
#submission Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#cross-references submitted to the Brookhaven Protein Data Bank, October 1996  
#journal PDB:1VJS  
#book TN032244  
#title in Enzymes for Carbohydrate Engineering (In: Prog. Biotechnol., V.12), Park, K.H.  
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#title Crystallization and a preliminary X-ray crystallographic study of alpha-amylase from bacillus licheniformis.  
COMMENT Resolution: 1.7 angstroms  
COMMENT Determination: X-ray diffraction  
KEYWORDS R-value: 0.199  
glycosyltransferase; glycosyltransferase; hydrolase  
FEATURE  
19-25 #region helix (right hand alpha)\  
27-32 #region helix (right hand alpha)\  
78-91 #region helix (right hand alpha)\  
155-157 #region helix (right hand alpha)\  
5-7,37-39,94-99 #region beta sheet\  
173-175,109-117, #region beta sheet\  
131-138 #region beta sheet\  
158-161 #region beta sheet\  
SUMMARY #length 179 #checksum 2012

Query Match 26.0%; Score 1006; DB 5; Length 179;  
Best Local Similarity 70.9%; Pred. No. 6.06e-156;  
Matches 127; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

Db 2 NGTLMQYFEWTPNDGQHWKRLQNDAAANLSKSGITAVWIPPAWKGTSONDVGAYDLYD 61  
QY 37 NGTMMQYFEWHPNDGQHWKRLQNDAAANLSKSGITAVWIPPAWKGTSONDVGAYDLYD 96  
Db 62 LGEPHQKGTVRTKYGKELQSAIKSLHSRDIINVYGDVVVNNHKGADGTEMVNAVEVDP 121  
QY 97 LGEPHQKGTVRTKYGKELQSAIKSLHSRDIINVYGDVVVNNHKGADGTEMVNAVEVNR 156  
Db 122 DNRNVISGEHLIKANTHFFPGRGSTYSDFKWHWHYHFDGTDWDESRLK-NRIYKFOGKA 179  
QY 157 NRNOEISGEYTIETAWTKEDFPGRGNTHSNFKRWYHFDGTDWDSRLQNLKIKYKFRGTG 215

RESULT 15  
ENTRY TITLE  
ALTERNATE\_NAMES 1A1506 #type fragments  
PDB\_TITLE alpha-amylase (EC 3.2.1.1) - Bacillus amyloliquefaciens  
ORGANISM (fragments)  
#formal\_name Bacillus amyloliquefaciens

DATE 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change  
31-Oct-1997  
ACCESSIONS A91760; A91759; A19506  
REFERENCE A91760  
#authors Sachdev, O.; Friedberg, F.  
#journal Int. J. Pept. Protein Res. (1981) 18:228-236  
#cross-references MUID:82189140  
#accession A91760  
#molecule\_type protein  
#residues 1-56; 57-144 #label SAC  
##note residues 1-56 correspond to residues 229-287 of the  
complete sequence; residues 57-144 (without residues  
140-144) correspond to residues 335-397 of the  
complete sequence

REFERENCE A91759  
#authors Detera, S.D.; Friedberg, F.  
#journal Int. J. Pept. Protein Res. (1981) 17:93-106  
#cross-references MUID:81191186  
#accession A91759  
#molecule\_type protein  
#residues 145-217 #label DET  
##note this sequence corresponds to residues 398-469 of the  
complete sequence

COMMENT See ALBSN.  
FUNCTION  
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic  
bonds  
#pathway glycogen/starch degradation  
CLASSIFICATION #superfamily alpha-amylase, amyloliquefaciens type;  
#alpha-amylase core homology  
KEYWORDS glycosidase; hydrolase; polysaccharide degradation  
SUMMARY #length 217 #checksum 7645

Query Match 8.9%; Score 344; DB 2; Length 217;  
Best Local Similarity 60.8%; Pred. No. 1.87e-36;  
Matches 59; Conservative 14; Mismatches 13; Indels 11; Gaps 7;  
Db 118 AYWAFILTRREGYPQVFGDVGSGVMYGTGTPKPEISLKDNIPIKARKEYAYGPO 177  
QY 383 AY-ALILTRQO-GYPSVFGD---Y-YGI-PT--HGVPFMSKSIDPLQARQTYAIGTQ 432  
Db 178 HDYIDPH-VIGWTRGDSAAKSLAALISDGPGGK 213  
QY 433 HDYFDHHDIIIGWTRGDSHPNSGLATIMSDGPGNK 469

Search completed: Tue Sep 15 14:17:58 1998  
Job time : 81 secs.

\*\*\*\*\*  
W P S R L H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:13:48 1998; MasPar time 22.56 Seconds  
Tabular output not generated. 370.088 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHRIISVLLTLLAVAV.....ADCGWNTYNGGAVSVWVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq32  
1:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 35.739; Variance 162.847; scale 0.219  
pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match %	Length DB ID	Description Pred. No.
1	3873	100.0	516 24	W11326 Alkaline liquefying a 0.00e+00
2	3539	91.4	485 27	W31499 Bacillus sp. alpha am 0.00e+00
3	3537	91.3	485 20	W12131 Alpha-amylase variant 0.00e+00
4	3536	91.3	485 20	W12120 Alpha-amylase variant 0.00e+00
5	3536	91.3	485 20	W12125 Alpha-amylase variant 0.00e+00
6	3536	91.3	485 20	W12128 Alpha-amylase variant 0.00e+00
7	3536	91.3	485 20	W12118 Alpha-amylase variant 0.00e+00
8	3536	91.3	485 20	W12122 Alpha-amylase variant 0.00e+00
9	3536	91.3	485 20	W12123 Alpha-amylase variant 0.00e+00
10	3536	91.3	485 20	W12130 Alpha-amylase variant 0.00e+00
11	3536	91.3	485 20	W12129 Alpha-amylase variant 0.00e+00
12	3536	91.3	485 20	W12127 Alpha-amylase variant 0.00e+00
13	3536	91.3	485 20	W12126 Alpha-amylase variant 0.00e+00
14	3536	91.3	485 20	W12124 Alpha-amylase variant 0.00e+00
15	3536	91.3	485 20	W12108 Alpha-amylase variant 0.00e+00
16	3536	91.3	485 20	W12119 Alpha-amylase variant 0.00e+00
17	3536	91.3	485 20	W12121 Alpha-amylase variant 0.00e+00
18	3534	91.2	485 20	W12114 Alpha-amylase variant 0.00e+00

19	3534	91.2	485 20	W12117 Alpha-amylase variant 0.00e+00
20	3533	91.2	485 20	W12113 Alpha-amylase variant 0.00e+00
21	3532	91.2	485 20	W12105 Alpha-amylase variant 0.00e+00
22	3532	91.2	485 20	W12099 Alpha-amylase variant 0.00e+00
23	3532	91.2	485 20	W12106 Alpha-amylase variant 0.00e+00
24	3532	91.2	485 20	W12100 Alpha-amylase variant 0.00e+00
25	3532	91.2	485 20	W12098 Alpha-amylase variant 0.00e+00
26	3532	91.2	485 20	W12104 Alpha-amylase variant 0.00e+00
27	3532	91.2	485 20	W12102 Alpha-amylase variant 0.00e+00
28	3532	91.2	485 20	W12107 Alpha-amylase variant 0.00e+00
29	3532	91.2	485 20	W12103 Alpha-amylase variant 0.00e+00
30	3529	91.1	485 20	W12110 Alpha-amylase variant 0.00e+00
31	3528	91.1	485 20	W12109 Alpha-amylase variant 0.00e+00
32	3528	91.1	485 20	W12111 Alpha-amylase variant 0.00e+00
33	3528	91.1	485 20	W12144 Alpha-amylase variant 0.00e+00
34	3528	91.1	485 20	W12112 Alpha-amylase variant 0.00e+00
35	3528	91.1	485 20	W12101 Alpha-amylase variant 0.00e+00
36	3513	90.7	485 16	R81835 Bacillus sp. alkaline 0.00e+00
37	3505	90.5	483 20	W12116 Alpha-amylase variant 0.00e+00
38	3503	90.4	483 20	W12142 Alpha-amylase variant 0.00e+00
39	3503	90.4	483 20	W12141 Alpha-amylase variant 0.00e+00
40	3501	90.4	483 20	W12115 Alpha-amylase variant 0.00e+00
41	3503	90.4	485 20	W12132 Alpha-amylase variant 0.00e+00
42	3495	90.2	483 20	W12133 Alpha-amylase variant 0.00e+00
43	3489	89.6	483 20	W12134 Alpha-amylase variant 0.00e+00
44	3464	89.4	483 20	W12135 Alpha-amylase variant 0.00e+00
45	3453	89.2	483 20	W12136 Alpha-amylase variant 0.00e+00

ALIGNMENTS

RESULT 1  
ID W11326 standard; Protein; 516 AA.  
AC W11326;  
DT 11-NOV-1997 (first entry)  
DE Alkaline liquefying alpha-amylase.  
KW Alkaline liquefying alpha-amylase; Bacillus; alpha-amylase; detergent;  
KW starch-related polysaccharide; hydrolysis; enzyme; surfactant resistance;  
KW alpha-1,4-glucosidic link; alkalophilic Bacillus; laundry detergent;  
OS dish-washing detergent; starch.  
OS Bacillus species KSM-AP1378.  
PN WO9700324-A1.  
PD 03-JAN-1997.  
PF 14-JUN-1986; J01641.  
PR 14-JUN-1995; JP-147257.  
PA (KAOS ) KAO CORP.  
PI Ara K, Hatada Y, Ito S, Kawai S, Ozaki K;  
DR WPI; 97-118708/11.  
DR N-PSDB; T51339.  
PT DNA encoding alkaline liquefying alpha-amylase - useful in  
PT dish-washing and laundry detergents for removal of starch dirt  
PS Claim 2: Page 23-26; 40pp; English.  
CC This sequence represents an alkaline liquefying alpha-amylase.  
CC Alpha-amylase is an enzyme that acts on starch-related polysaccharides,  
CC hydrolysing the alpha-1,4-glucoside bond of the polysaccharide molecule.  
CC Alkaline liquefying alpha-amylases exhibit resistance to surfactants used  
CC in detergents, and decompose starch or starch-related polysaccharides in  
CC a highly random manner. The Bacillus species KSM-AP1378, from which this  
CC sequence was isolated, is an alkalophilic Bacillus strain. It was  
CC isolated from soil in the vicinity of the city of Tochigi. The enzyme is  
CC useful in improving the efficiency of dish-washing and laundry  
CC detergents, particularly on starch dirt.  
SQ Sequence 516 AA;

Query Match 100.0%; Score 3873; DB 24; Length 516;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 mklhriisvlltlllavavlfpytmepaqahhngtngtmqyfewhlpdngnhwrlrd 60  
QY 1 MKLHRIISVLLTLLAVAVLFPPYTMPEAQAHNGTNGTMMQYFEWHLPDNGNHWRLLRD 60  
Db 61 daanlkskgtavwipawkgsqndvgaydyldlgefnkgvtvtrktytrslqgav 120

```

QY 61 DAANLKSIGITAVWIPPAWKTQSDVGYGAYDLYDLGEFNQKGTVRTKYGTRSQLQAV 120
Db 121 tslnknngiqvvgdvvmhkhgagadtemvnavesnrnsrqeaisgeyieawtkkfdfpgrg 180
QY 121 TSLKNNGIQVGVVMMHKGAGADTEMVNAVESNRNSRQEIISGEYIEAWTKFDFPGRG 180
Db 181 nthsnfkwrwhfdgtgdwdsgrlqnkiykfrgtgkdwedvdiengnydylmavadidmd 240
QY 181 NTHSNFKWRWHFDGTGDWDSRQLQNKIYKFRGTGKADWEVDIENGNYDYLMAVADIDMD 240
Db 241 hpevinealnrgwvyntlnldgfridavkhikysytrdwlthvtrnttggkpmfavaefwk 300
QY 241 HPEVINEALNRGWVYNTLNLDGFRIDAVKHIKYSYTRDWLTHVTRNTTGGKPMFAVAEFWK 300
Db 301 ndlaaenylnktswnhsvfdvplhynlnasnggyfdmrnlilngsvvqkphihavtfv 360
QY 301 NDLAATENYLNKTSWNHSHVFDVPLHYNLYNASNGGYFDMRNLILNGSVVQKPHIHAVTFV 360
Db 361 dnhdsqgealesfvgsfwkplayallltregqpsvfygdygipthgvpmskskidpl 420
QY 361 DNHDSQGEALESFVGSFWKPLAYALLLTREGQPSVFGDYGIPTHGVPMSKSIDPL 420
Db 421 lqarqiyavgtqhdvfdhddiqtregdshpnsqglatimsdpgpgnkmyvvgkhkag 480
QY 421 LQARQIYAVGTQHDYFDHDDIQTREGDSSHPNSGLATIMSDPGPGNKMYVVGKHKAG 480
Db 481 vwrldtgnrsgtvtinadgwnftvnggavsvvkw 516
QY 481 VWRDITGNRSGTVTINADGWNFTVNGGAVSVVWK 516

```

## RESULT 2

```

ID W31499 standard; protein; 485 AA.
AC W31499;
DT 08-APR-1998 (first entry)
DE Bacillus sp. alpha amylase.
KW Alpha amylase; hard surface cleaning; dishwashing; laundry.
OS Bacillus sp.
PN W09732361-A2.
PD 12-SEP-1997.
PF 04-MAR-1997; U03635.
PR 07-MAR-1996; WO-U03276.
PA (PROC ) PROCTER & GAMBLE CO.
PI Baek AC, Jones IA, Ohtani R, Pramod K, Rai S,
PI Showell MS, Ward G;
DI WPI; 97-457524/42.
PT Detergent compositions for hard surface cleaning and laundry use -
PT contains Bacillus derived alpha amylase with improved
PT thermostability, reduced calcium ion dependency etc.
PS Claim 1; Pages 86-87; 97pp; English.
CC The present sequence is a Bacillus sp. alpha amylase with a
CC specific activity at least 25% higher than that of Termamyl (RTM)
CC at 25 to 55 degrees C at pH 8 to 10, measured by the Phadebas (RTM)
CC activity test. It is of use in hard surface cleaning, hand or
CC machine dishwashing and laundry at a temperature of 40 to 25
CC degrees C. Improved cleaning, stain removal and fabric care are
CC obtained by using it at a concentration of 0.00018 to 0.06%
SQ Sequence 485 AA;

```

```

Query Match 91.4%; Score 3539; DB 27; Length 485;
Best Local Similarity 95.1%; Pred. No. 0.00e+00;
Matches 461; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

Db 1 hncgtngtmmygfewylpndgnhwnrirdaanklskgitavwppawkgtsqndvgyga 60
QY 32 HHNGTNGTMGYFEWHLPNDGNHWNRLRDDAANKLSKGITAVWPPAWKGTSDNDVGYGA 91
Db 61 ydlldlgefngkgvtrkygtrnglqaavtslknngiqvvgdvvmhkhgagdeivnav 120
QY 92 YDLIDLGEFNKGVTWKYGTGRSQLQAVTSLKNGIQVGVVMMHKGAGDTEMVNAV 151
Db 121 evnrsnrnqetsgeyaieawtkkfdfpgrgnhssfkwrwhfdgtgdwdsgrlqnkiyk 180

```

```

QY 152 EVNSRNRQELSGSYTTEAWTKFDFPGRGNTHSNFKWRWHYFDGTDMDSRQLQNKIYKF 211
Db 181 rgtgkdwedvtenpnydylmavadidmdhpevihealnrgwvyntlnldgfridavkh 240
QY 212 RGTGKADWEVDIENGNYDYLMAVADIDMDHPEVINEALNRGWVYNTLNLDGFRIDAVKH 271
Db 241 ikysytrdwlthvtrnttggkpmfavaefwkndlgaienylnktswnhsvfdvplhynlna 300
QY 272 IKYSYTRDWLTHVTRNTTGGKPMFAVAEFWKNDLGAIAENYLNKTSWNHSHVFDVPLHYNLYNA 331
Db 301 snsggyfdmrnlilngsvvqkphthavtfvndhdsqgealesfvgsfwkplayallltre 360
QY 332 SNSGGYFDMRNLILNGSVVQKPHTHAVTFVNDHDSQGEALESFVGSFWKPLAYALLTRE 391
Db 361 qgypsvfygdygipthgvpmskskidpllqarqiyavgtqhdvfdhddiqtregdss 420
QY 392 QGYPSVFGDYGIPTHGVPMSKSIDPLLQARQIYAVGTQHDYFDHDDIQTREGDSS 451
Db 421 hpnsglatimsdpgpgnkmyvvgkhkagvwrldtgnrtgtvtinadgwnftvnggavsv 480
QY 452 HPNSGLATIMSDPGPGNKMYVVGKHKAGVWRDITGNRSGTVTINADGWNFTVNGGAVS 511
Db 481 vvwkq 485
QY 512 VVWKQ 516

```

## RESULT 3

```

ID W12131 standard; protein; 485 AA.
AC W12131;
DT 08-APR-1997 (first entry)
DE Alpha-amylase variant N106D.
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;
KW calcium ion dependency; alpha-amylolytic activity; washing composition;
KW textile desizing; papermaking; beer-making; ethanol production;
KW sweetener.
OS Synthetic.
FH Key
FT misc_difference 106
FT /label= N106D
PN W09623873-A1.
PD 08-AUG-1996.
PF 05-FEB-1996; DK0056.
PR 03-FEB-1995; DK-000126.
PR 29-MAR-1995; DK-000336.
PR 29-SEP-1995; DK-001097.
PR 06-OCT-1995; DK-001121.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgaard-frantzen H, Borchert T, Svendsen A;
DI WPI; 96-371423/37.
PT Alpha-amylase variants - with improved thermal and oxidation
PT stability and reduced calcium ion dependency
PS Claim 11; ; ilpp; English.
CC W12098-W12144 represent alpha-amylase variants of the invention. The
CC variants of the invention were created using site directed, or random,
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,
CC W12142 and W12144 are specifically variants of the alkaphilic bacillus
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have
CC improved thermal stability (such as at temperatures in the range of 40-70
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion
CC dependency. The variants can also have increased alpha-amylolytic
CC activity (especially at pH values in the range of 8.5-10.5), and improved
CC binding of a particular substrate. These variant alpha-amylases also
CC possess improved specificity to a particular substrate, and/or improved
CC specificity with respect to cleavage of substrate. These sequences can
CC be used in detergent and washing compositions, and for textile desizing.
CC The alpha-amylase variants can also be used in papermaking and
CC beer-making processes. These variants can also be used in the production
CC of sweeteners and ethanol from starch.
SQ Sequence 485 AA;

```

Query Match 91.3%; Score 3537; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. NO. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtngmmgyfeywlpndgnhwnrlrddaanlkskgtavwipawkgtsqndvggca 60  
QY 32 HHNGTNGTMMGYFEWHLNDGNHWNRLRDDAANLKSIGTAVWIPPAWKGTSQNDVGCGA 91

Db 61 ydlvdlgefngkgtvrtkygtrnqlqaavtslknggiqyvgdvnmhkggagdtseivnav 120  
QY 92 YDLVDLGEFNGKGTVRTKYGTRSQLQGAVTSLKNGGIQYVGDVNMHKGAGDGTSEMVNAV 151

Db 121 evnsrnrnqetgsgeyaieawtkfdpgrgnhssfkwrwyhfdgtdwdsrqlqknkiyxf 180  
QY 152 EVNSRNRNQEISGEYIEAWTKFDPGRGNTHSNFKWRWYHFDGTDWDSRQLQKNKIYKF 211

Db 181 rgtgkdwewdvtengnydylmayadvmdhpevihelrnwgvwyntnlldgfridavkh 240  
QY 212 RGTGKADWEVDIENGNDYLMYADIDMDHPEVIELRNWGVWYNTNLNLDGFRIDAVKH 271

Db 241 ikysfrdwlthvrnttkpmafavaefwkndlgaienylnktswnhsvfdvplhynyna 300  
QY 272 IKYSYTRDWLTHVRNTTKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYNYNA 331

Db 301 snsggyydmrnlngsvvqkhpthavtfvndhdsqgealesfvqgwkfplayalvtre 360  
QY 332 SNSGGYFDMRNLNGSVVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391

Db 361 qgypsvfygdygipthgvpamkskidpqlqarqtfaygtghdyfdhddiigwtregss 420  
QY 392 QGYPSVFGDYGIPTHGVPMSKSIDPQLQARQTYAYGTQHDYFDHDDIIGWTREGDSS 451

Db 421 hpnsglatimsdpggnkwmvvgkknagqwrdrtdgnrtgtvtinadgwnfsvnggsvs 480  
QY 452 HPNSGLATIMSDPGGNKWMVVGKKNAGQWRDRITGNRSSTVTINADGWNFTVNGGAVS 511

Db 481 vvwkq 485  
QY 512 VVWQK 516

RESULT 4  
ID W12120 standard; protein; 485 AA.  
AC W12120;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant K239R.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 239 /label= K239R  
FT FT  
FN W09623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1995; DK0056.  
PR 03-FEB-1995; DK-000336.  
PR 29-MAR-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PR (NOVO) NOVO-NORDISK AS.  
PI Bisgaard-frantzen H, Borchert T, Svendsen A;  
DR WPI: 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Claim 11; : ilpp: English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have

CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrates, and/or improved  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. NO. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtngmmgyfeywlpndgnhwnrlrddaanlkskgtavwipawkgtsqndvggca 60  
QY 32 HHNGTNGTMMGYFEWHLNDGNHWNRLRDDAANLKSIGTAVWIPPAWKGTSQNDVGCGA 91

Db 61 ydlvdlgefngkgtvrtkygtrnqlqaavtslknggiqyvgdvnmhkggagdtseivnav 120  
QY 92 YDLVDLGEFNGKGTVRTKYGTRSQLQGAVTSLKNGGIQYVGDVNMHKGAGDGTSEMVNAV 151

Db 121 evnsrnrnqetgsgeyaieawtkfdpgrgnhssfkwrwyhfdgtdwdsrqlqknkiyxf 180  
QY 152 EVNSRNRNQEISGEYIEAWTKFDPGRGNTHSNFKWRWYHFDGTDWDSRQLQKNKIYKF 211

Db 181 rgtgkdwewdvtengnydylmayadvmdhpevihelrnwgvwyntnlldgfridavrh 240  
QY 212 RGTGKADWEVDIENGNDYLMYADIDMDHPEVIELRNWGVWYNTNLNLDGFRIDAVKH 271

Db 241 ikysfrdwlthvrnttkpmafavaefwkndlgaienylnktswnhsvfdvplhynyna 300  
QY 272 IKYSYTRDWLTHVRNTTKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYNYNA 331

Db 301 snsggyydmrnlngsvvqkhpthavtfvndhdsqgealesfvqgwkfplayalvtre 360  
QY 332 SNSGGYFDMRNLNGSVVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391

Db 361 qgypsvfygdygipthgvpamkskidpqlqarqtfaygtghdyfdhddiigwtregss 420  
QY 392 QGYPSVFGDYGIPTHGVPMSKSIDPQLQARQTYAYGTQHDYFDHDDIIGWTREGDSS 451

Db 421 hpnsglatimsdpggnkwmvvgkknagqwrdrtdgnrtgtvtinadgwnfsvnggsvs 480  
QY 452 HPNSGLATIMSDPGGNKWMVVGKKNAGQWRDRITGNRSSTVTINADGWNFTVNGGAVS 511

Db 481 vvwkq 485  
QY 512 VVWQK 516

RESULT 5  
ID W12125 standard; protein; 485 AA.  
AC W12125;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant D199N.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 199 /label= D199N  
FT FT  
FN W09623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1995; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.

8/8/96  
not published

PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PS stability and reduced calcium ion dependency  
PT Claim 11: 111pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkalophilic bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
Db 1 hngtngtmqyfwylpndgnhwnlrddaanlkskgitavwippawkgtsqndvgyga 60  
QY 32 HHNGTNGTMQYFEWHLPGNDGNHWNRLRDDAANLKSIGITAVWIPPAAWKGTSQNDVGYGA 91  
Db 61 ydlldlgefnqkgtrtkygrtnqlqaavtslknngiqvgydvvmhkggadgteivnav 120  
QY 92 YDLDLGEFNGKGTVRTKYGRSLOQAVTSLKNNGIQVYGDVVMHKGADGTEMVNAV 151  
Db 121 evnrsnrngetsgeyaieawtkdfpgrgnhshsfkwrwyhfdgtdwdqrqlqknykf 180  
QY 152 EVNRSNRNGEISGTYTEAWTKDFPGRGNTHSNFKRWYHFDGTDWDQSRQLQNKIYKF 211  
Db 181 rgtgkawdwevdtengnynlymadvdmhpevihelrnwgvvtyntlnldgfridavkh 240  
QY 212 RGTGKAWDWEVDIENGNDYLMYADIDMDHPEVINELRNKGWVWYTNLNDGFRIDAVKH 271  
Db 241 ikysfrdwlthvrnttgkpmfavaefwknldgalenylnktswnhsfvdpplhynlyna 300  
QY 272 IKYSYTRDWLTHVRNTTGKPMFAVEFWKNDLAAIENYLNKTSWNHSFVDPPLHYNLYNA 331  
Db 301 snsggyvdmrnlilngsvvqkphthavtfdvndhdsqgealesfvqgwfplayalvtre 360  
QY 332 SNSGGYFDMRNLILNGSVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391  
Db 361 qgypsvfydygipthgvpamkskidpllarqqtfaigtqhdvfdhhdliigwtregns 420  
QY 392 QGYPSPVFGDYIGIPTHGVPMSKSIDPLLARQQTAYGTQHDYFDHHDIIIGWTREGSDS 451  
Db 421 hpnsglatimsdpggknkmyvgknkagqvrtditgnrtgtvtinadgwnfsvnggsvs 480  
QY 452 HPNSGLATIMSDPGGKNKMYVGKHKAGQVWRDITGNRSGTGTVINADGWNFTVNGGAVS 511  
Db 481 vvvkq 485  
QY 512 VVWKQ 516

RESULT 6

ID W12128 standard; protein; 485 AA.  
AC W12128;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant D209N  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;

KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 209 /label= D209N  
FT W09623873-Al.  
PN 08-AUG-1996.  
PD 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Claim 11: 111pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkalophilic bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
Db 1 hngtngtmqyfwylpndgnhwnlrddaanlkskgitavwippawkgtsqndvgyga 60  
QY 32 HHNGTNGTMQYFEWHLPGNDGNHWNRLRDDAANLKSIGITAVWIPPAAWKGTSQNDVGYGA 91  
Db 61 ydlldlgefnqkgtrtkygrtnqlqaavtslknngiqvgydvvmhkggadgteivnav 120  
QY 92 YDLDLGEFNGKGTVRTKYGRSLOQAVTSLKNNGIQVYGDVVMHKGADGTEMVNAV 151  
Db 121 evnrsnrngetsgeyaieawtkdfpgrgnhshsfkwrwyhfdgtdwdqrqlqknykf 180  
QY 152 EVNRSNRNGEISGTYTEAWTKDFPGRGNTHSNFKRWYHFDGTDWDQSRQLQNKIYKF 211  
Db 181 rgtgkawdwevdtengnynlymadvdmhpevihelrnwgvvtyntlnldgfridavkh 240  
QY 212 RGTGKAWDWEVDIENGNDYLMYADIDMDHPEVINELRNKGWVWYTNLNDGFRIDAVKH 271  
Db 241 ikysfrdwlthvrnttgkpmfavaefwknldgalenylnktswnhsfvdpplhynlyna 300  
QY 272 IKYSYTRDWLTHVRNTTGKPMFAVEFWKNDLAAIENYLNKTSWNHSFVDPPLHYNLYNA 331  
Db 301 snsggyvdmrnlilngsvvqkphthavtfdvndhdsqgealesfvqgwfplayalvtre 360  
QY 332 SNSGGYFDMRNLILNGSVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391  
Db 361 qgypsvfydygipthgvpamkskidpllarqqtfaigtqhdvfdhhdliigwtregns 420  
QY 392 QGYPSPVFGDYIGIPTHGVPMSKSIDPLLARQQTAYGTQHDYFDHHDIIIGWTREGSDS 451  
Db 421 hpnsglatimsdpggknkmyvgknkagqvrtditgnrtgtvtinadgwnfsvnggsvs 480

```

QY 452 HPNSGLATMSDGPGGKWMYVGHKAGQVWRDITGNRSSTVTINADGNGFTVNGGAVS 511
Db 481 vvvkq 485
QY 512 VVWKQ 516

RESULT 7
ID W12118 standard; protein; 485 AA.
AC W12118;
DT 08-APR-1997 (first entry)
DE Alpha-amylase variant K108R
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;
KW calcium ion dependency; alpha-amylolytic activity; washing composition;
KW textile desizing; papermaking; beer-making; ethanol production;
OS Sweetener.
QY Synthetic.
FH Key
FT misc_difference 108
FT /label= K108R
PN WO9623873-A1.
PD 08-AUG-1996.
PF 05-FEB-1996; DK0056.
PR 03-FEB-1995; DK-000126.
PR 29-MAR-1995; DK-000336.
PR 29-SEP-1995; DK-001097.
PR 06-OCT-1995; DK-001121.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgard-frantzen H, Borchert T, Svendsen A;
DR WPI; 96-371423/37.
PT Alpha-amylase variants - with improved thermal and oxidation
PT stability and reduced calcium ion dependency
PS Claim 11; 11lpp; English.
CC W12098-W12144 represent alpha-amylase variants of the invention. The
CC variants of the invention were created using site directed, or random,
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have
CC improved thermal stability (such as at temperatures in the range of 40-70
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion
CC dependency. The variants can also have increased alpha-amylolytic
CC activity (especially at pH values in the range of 8.5-10.5), and improved
CC binding of a particular substrate. These variant alpha-amylases also
CC possess improved specificity to a particular substrate, and/or improved
CC specificity with respect to cleavage of substrate. These sequences can
CC be used in detergent and washing compositions, and for textile desizing.
CC The alpha-amylase variants can also be used in papermaking and
CC beer-making processes. These variants can also be used in the production
CC of sweeteners and ethanol from starch.
SQ Sequence 485 AA;

Query Match
Best Local Similarity 91.3%; Score 3536; DB 20; Length 485;
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hngtngtmqyfeywlpndgnhwnlrddaanlkskgitavwippawkgtsqndvggga 60
QY 32 HNGTNGTMMQYFEWLPNDGNHWNLRDDAANLKSKITAVWIPPAAWKTSQNDVGGA 91

Db 61 ydlldlgefnqgvtvtygtrnqlqaavtslknnglqvgydvmmhrggadgteivnav 120
QY 92 YDLIDLGEFNQGTVRTKYGTRSLQAGVTSKNNGLQVGYDVMMHKGADGTEMVNAV 151

Db 121 evnrsnrngetageyaieawtkdfpgrgnhshfkrwyghfdgtdwdqgrlqnkykf 180
QY 152 EVNRSNRNEISGEYIEAWTKDFPGRGNTHSNFKRWYHFDGTDWDQGRQLQNKYKF 211

Db 181 rgtgkdwdevdtenqydylymadvdmhdpevihelrnwgvvytntlnldgfridavkh 240
QY 212 RTGKADWDEVDIENQYDLYMADVDMHDPEVINELRNMGVWYTYTLNLDGFRIDAVKH 271

Db 241 ikysfrtdwlthvtrnttgkpmfavaefwkndlgaienylnktswnhsvfdplhynlna 300

```

```

QY 272 IKSYTRDMLTHVTRNTTGKPMFAVAEFWKNDLAAIENTLNKTSWNHSVDFDPLHYNLA 331
Db 301 snsggyvdmrnlpgsvvqkhphtbavtfvdhdsqpgaelesfvqgwkfpalayalvtre 360
QY 332 SNSGGYFDMRNLNGSVVQKHPIHAVTFVDHDSQPGAELESFVQSWFKPLAYALITRE 391
Db 361 qgvpsvfgydygigphthgvpmkskidlqarqtfaygtqhdhdiigwtregss 420
QY 392 QGVPSVFYGYGIGPIHTHGVPSMKSKIDPLQARQTFAYGTQHDYDHDHDIIGWTREGDSS 451
Db 421 hpnsglatmsdgpoggknwmyvghkagqvwrditgnrtgtvtinadgngfsvnggsvs 480
QY 452 HPNSGLATMSDGPGGKWMYVGHKAGQVWRDITGNRSSTVTINADGNGFTVNGGAVS 511
Db 481 vvvkq 485
QY 512 VVWKQ 516

RESULT 8
ID W12122 standard; protein; 485 AA.
AC W12122;
DT 08-APR-1997 (first entry)
DE Alpha-amylase variant D163N.
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;
KW calcium ion dependency; alpha-amylolytic activity; washing composition;
KW textile desizing; papermaking; beer-making; ethanol production;
KW Sweetener.
OS Synthetic.
FH Key
FT misc_difference 163
FT /label= D163N
PN WO9623873-A1.
PD 08-AUG-1996.
PF 03-FEB-1996; DK0056.
PR 03-FEB-1995; DK-000126.
PR 29-MAR-1995; DK-000336.
PR 29-SEP-1995; DK-001097.
PR 06-OCT-1995; DK-001121.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgard-frantzen H, Borchert T, Svendsen A;
DR WPI; 96-371423/37.
PT Alpha-amylase variants - with improved thermal and oxidation
PT stability and reduced calcium ion dependency
PS Claim 11; 11lpp; English.
CC W12098-W12144 represent alpha-amylase variants of the invention. The
CC variants of the invention were created using site directed, or random,
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have
CC improved thermal stability (such as at temperatures in the range of 40-70
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion
CC dependency. The variants can also have increased alpha-amylolytic
CC activity (especially at pH values in the range of 8.5-10.5), and improved
CC binding of a particular substrate. These variant alpha-amylases also
CC possess improved specificity to a particular substrate, and/or improved
CC specificity with respect to cleavage of substrate. These sequences can
CC be used in detergent and washing compositions, and for textile desizing.
CC The alpha-amylase variants can also be used in papermaking and
CC beer-making processes. These variants can also be used in the production
CC of sweeteners and ethanol from starch.
SQ Sequence 485 AA;

Query Match
Best Local Similarity 91.3%; Score 3536; DB 20; Length 485;
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hngtngtmqyfeywlpndgnhwnlrddaanlkskgitavwippawkgtsqndvggga 60
QY 32 HNGTNGTMMQYFEWLPNDGNHWNLRDDAANLKSKITAVWIPPAAWKTSQNDVGGA 91

```

Db 61 ydlvdlgefnkgkgtvrtkygtrnqlqaavtslknngiqvgydvmmhkggagdtelnav 120  
 Qy 92 YDLVDLGEFNGKGTVRTKYGTRNQLQAAVTSLKNNGIQVGYDVMMHKGAGDGTENAV 151  
 Db 121 evnsnqnqetsgevaaleawtkkfppgrgnhshsfkwyhfnngtdwqsrqlqnkiyxf 180  
 Qy 152 EVNSNRNQEISGEYTIKFTKDFPGRGNTHSNFKRWYHFDGTDWQSRQLQNKIYKF 211  
 Db 181 rgtgkavdwedvtengnydylmyadvmdhpevihelrnwgyvwtntlnldgfridavkh 240  
 Qy 212 RGTGKAWDEWDIENGNYDYLMYADIDMDHPEVINELRNWGYWTNTLNLDGFRIDAVKH 271  
 Db 241 ikysftrdwthvnttckgmfaevafwkndlgaienylnktswnhsvfdvplhynlyna 300  
 Qy 272 IKYSYTRDWLTHVNTTCKGMPFAEAFWKNDLAAIENLYNKTSWNHSHVDFVPLHYNLYNA 331  
 Db 301 snsggyydmnrlngsvvqkqhthavtfvdnhdsgpgealesfvqgwfplayalvtre 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHAVTFVDNHDSPGGEALSFVQSWFKPLAYALITRE 391  
 Db 361 qgypsvfgydygipthgvpamkskidpllqarqtavgtqhdvfdhhdliigtregns 420  
 Qy 392 QGYPSVFYGYDYGIPTHGVPAMSKIDPLLQARQTYAGTQHDYFDHHDIIIGWREGDSS 451  
 Db 421 hpnsglatinsdpggknkmyvgnkagqvwrtditgnrtgvtctinatdgwgnfsvngs 480  
 Qy 452 HPNSGLATINS DPGGKNKMYVGNKAGQVWRDITGNRSGTGTINADGWGNFTVNGGAVS 511  
 Db 481 vvwkq 485  
 Qy 512 VVWKQ 516

## RESULT 9

ID W12123 standard; protein; 485 AA.  
 AC W12123;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D188N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key  
 FT misc\_difference 188  
 FT /label= D188N  
 PN WO9623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; ililpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.

CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtngtmgyfeywlpndgnhwnrlrddaanlkskgtavwippawkgtsqndvggga 60  
 Qy 32 HHNGTNGTMMGYFEWHLPNNDGNHWNRLRDDAANLKS KGTAVWIPPAPWKGTSQNDVGGA 91  
 Db 61 ydlvdlgefnkgkgtvrtkygtrnqlqaavtslknngiqvgydvmmhkggagdtelnav 120  
 Qy 92 YDLVDLGEFNGKGTVRTKYGTRNQLQAAVTSLKNNGIQVGYDVMMHKGAGDGTENAV 151  
 Db 121 evnsnqnqetsgevaaleawtkkfppgrgnhshsfkwyhfdgtdwqsrqlqnkiyxf 180  
 Qy 152 EVNSNRNQEISGEYTIKFTKDFPGRGNTHSNFKRWYHFDGTDWQSRQLQNKIYKF 211  
 Db 181 rgtgkavdwedvtengnydylmyadvmdhpevihelrnwgyvwtntlnldgfridavkh 240  
 Qy 212 RGTGKAWDEWDIENGNYDYLMYADIDMDHPEVINELRNWGYWTNTLNLDGFRIDAVKH 271  
 Db 241 ikysftrdwthvnttckgmfaevafwkndlgaienylnktswnhsvfdvplhynlyna 300  
 Qy 272 IKYSYTRDWLTHVNTTCKGMPFAEAFWKNDLAAIENLYNKTSWNHSHVDFVPLHYNLYNA 331  
 Db 301 snsggyydmnrlngsvvqkqhthavtfvdnhdsgpgealesfvqgwfplayalvtre 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHAVTFVDNHDSPGGEALSFVQSWFKPLAYALITRE 391  
 Db 361 qgypsvfgydygipthgvpamkskidpllqarqtavgtqhdvfdhhdliigtregns 420  
 Qy 392 QGYPSVFYGYDYGIPTHGVPAMSKIDPLLQARQTYAGTQHDYFDHHDIIIGWREGDSS 451  
 Db 421 hpnsglatinsdpggknkmyvgnkagqvwrtditgnrtgvtctinatdgwgnfsvngs 480  
 Qy 452 HPNSGLATINS DPGGKNKMYVGNKAGQVWRDITGNRSGTGTINADGWGNFTVNGGAVS 511  
 Db 481 vvwkq 485  
 Qy 512 VVWKQ 516

## RESULT 10

ID W12130 standard; protein; 485 AA.  
 AC W12130;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant E194Q.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key  
 FT misc\_difference 194  
 FT /label= E194Q  
 PN WO9623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; ililpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The



CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12957, W12958, W12959, W12960, W12961,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 CC Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtmmqyfevylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvggga 60  
 Qy 32 HHNGTNGTMQYFEVHLPLNDGNHWNRLRDDAANLKSKITAVWIPPWAKGTSONDVGGA 91  
 Db 61 ydlvdlgefngqgtvrtkygttrnqlqaavtslknngiqyvgdvmmhkgagdtelnav 120  
 Qy 92 YDLVLDLGEFNQGTVRTKYGTTRSQLQAVTSLKNGIQYVGDVMMHKGAGDTENVNAV 151  
 Db 121 evnrsnrnqetsgeyaieawtkfdpgrgnhssfkwrwyhfdgtwdqsrqlqknykf 180  
 Qy 152 EVNRSNRNQETSGEYATTEAWTKFDPGRGNTHSNFKRWYHFDGTDWDSQROLQNKIYKF 211  
 Db 181 rgtgkwdvdtqngnydylmyadvmdhpevihelrnwgvwytnlndgfridavkh 240  
 Qy 212 RGTGKAWDVEVDIENGNYDYLMDYADIDMDHPEVINELRNWGVWYTNLNDGFRIDAVKH 271  
 Db 241 ikysftrdwlthvrttqkpmfavaefwkndlgaienylnktswnhsvfdvplhynlna 300  
 Qy 272 IKYSTROWLTHVRTTQKPMFAVEAFWKNDLAAIENYLNKTSWNHSDVDPVPLHYNLNA 331  
 Db 301 snsggyydmrnlngsvvqkphthavtfdvndhsqpgealesfvqqwfkplayalvltre 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHAVTFVDNHDSDPGEALSFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpqlqartfaygtqhdhdiigwtregss 420  
 Qy 392 QGYPSVFYGDYGIPTGHPVSMKSIDPQLQARTYAYGTQHDYFDHDIIGWTREGSS 451  
 Db 421 hpsnglatmsdpgggnkwmvvgknkagqvwrditgnrtgtvtinadgwnfsvnggsvs 480  
 Qy 452 HPNSGLATIMSDPGGNKWMYVGKHKAGQVWRDITGNRSGTVTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 Qy 512 VVWKQ 516

RESULT 11  
 ID W12129 standard; protein; 485 AA.  
 AC W12129;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant E190Q.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 190  
 FT /label= E190Q  
 PN W09623873-A1.

PD 08-AUG-1996.  
 PF 05-FEB-1996; DK00036.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; 11pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12957, W12958, W12959, W12960, W12961,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 CC Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtmmqyfevylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvggga 60  
 Qy 32 HHNGTNGTMQYFEVHLPLNDGNHWNRLRDDAANLKSKITAVWIPPWAKGTSONDVGGA 91  
 Db 61 ydlvdlgefngqgtvrtkygttrnqlqaavtslknngiqyvgdvmmhkgagdtelnav 120  
 Qy 92 YDLVLDLGEFNQGTVRTKYGTTRSQLQAVTSLKNGIQYVGDVMMHKGAGDTENVNAV 151  
 Db 121 evnrsnrnqetsgeyaieawtkfdpgrgnhssfkwrwyhfdgtwdqsrqlqknykf 180  
 Qy 152 EVNRSNRNQETSGEYATTEAWTKFDPGRGNTHSNFKRWYHFDGTDWDSQROLQNKIYKF 211  
 Db 181 rgtgkwdvdtqngnydylmyadvmdhpevihelrnwgvwytnlndgfridavkh 240  
 Qy 212 RGTGKAWDVEVDIENGNYDYLMDYADIDMDHPEVINELRNWGVWYTNLNDGFRIDAVKH 271  
 Db 241 ikysftrdwlthvrttqkpmfavaefwkndlgaienylnktswnhsvfdvplhynlna 300  
 Qy 272 IKYSTROWLTHVRTTQKPMFAVEAFWKNDLAAIENYLNKTSWNHSDVDPVPLHYNLNA 331  
 Db 301 snsggyydmrnlngsvvqkphthavtfdvndhsqpgealesfvqqwfkplayalvltre 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHAVTFVDNHDSDPGEALSFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpqlqartfaygtqhdhdiigwtregss 420  
 Qy 392 QGYPSVFYGDYGIPTGHPVSMKSIDPQLQARTYAYGTQHDYFDHDIIGWTREGSS 451  
 Db 421 hpsnglatmsdpgggnkwmvvgknkagqvwrditgnrtgtvtinadgwnfsvnggsvs 480  
 Qy 452 HPNSGLATIMSDPGGNKWMYVGKHKAGQVWRDITGNRSGTVTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 Qy 512 VVWKQ 516

RESULT 12

ID W12127 standard; protein; 485 AA.  
 AC W12127;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D207N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 207 /label= D207N  
 FT PN  
 PD W09623873-A1.  
 PF 08-AUG-1996.  
 PR 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrates. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 CC Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfevylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HHNGTNGTMMQYFEWHLPLNDGNHNRLRDDAANLKSIGITAVWIPPWKGTSQNDVGYGA 91  
 Db 61 ydlidgfeqndqgkvtvktygrnqlqaavtslknngiqvgydvvmnhkkgadgteivnav 120  
 QY 92 YDLIDLGEFNGKQVTRTKYGRSOLQGAVALSLKNGIQVGYDVVMNHKKGADGTEMVNAV 151  
 Db 121 evnrsnrgtsgyaleavtkdfpgrgnhshsfkwrwyhfdgtdwdqsrqlnkiyxf 180  
 QY 152 EVNRSNRNOEISGEYTTIATWTKDFPGRGNTSHSNFKRWYHFDGTDWDQSRQLNKIYKF 211  
 Db 181 rgtgkdwdevdteengnydylmadvnmhpeviheirngvvytntlnldgfridavkh 240  
 QY 212 RGTGKAWDDEVDTEENGYDYLMDYADIMDHPEVINELRNNGVWYNTNLNDGFRIDAVKH 271  
 Db 241 ikysfirdwlthvrnttqkpmfavaefwkdnlgaieynlnktswnhsvfdvplhnylna 300  
 QY 272 IKYSYTRDMLTHVRNTTQKPMFAVAEFAEWKNDLAAIENYLNKTSWNHVSFVDVPLHNYLNA 331  
 Db 301 snsggydmrnlngsvvqkphchavtfvndhdsqgpealesfvqqwfkplayalvltre 360  
 QY 332 SNSGGYFDMRNLNGSVVQKPHCHAVTFVNDHDSQGPQEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgyvpsfygdyggyipthgvpamkskldpllqarqtqfaygtqchdyfddhddliigtregss 420

QY 392 QGYPSVFYGYGIPTHGVPSMKSDPLQARQTVAYGTQHDYFDHDDIIGWTREGDSS 451  
 Db 421 hpsnlatimsdpggnkwmvgygkagvwdrtgnrtgtvtinadgwnfsvngsvs 480  
 QY 452 HPNSGLATIMSDPGGNKWMVGYGKAGQVWRDITGRSGTGTINADGWNFTVNGAVS 511  
 Db 481 vvwkq 485  
 QY 512 VVWKQ 516  
 RESULT 13  
 ID W12126 standard; protein; 485 AA.  
 AC W12126;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D205N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 205 /label= D205N  
 FT PN  
 PD W09623873-A1.  
 PF 08-AUG-1996.  
 PR 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrates. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 CC Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfevylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HHNGTNGTMMQYFEWHLPLNDGNHNRLRDDAANLKSIGITAVWIPPWKGTSQNDVGYGA 91  
 Db 61 ydlidgfeqndqgkvtvktygrnqlqaavtslknngiqvgydvvmnhkkgadgteivnav 120  
 QY 92 YDLIDLGEFNGKQVTRTKYGRSOLQGAVALSLKNGIQVGYDVVMNHKKGADGTEMVNAV 151  
 Db 121 evnrsnrgtsgyaleavtkdfpgrgnhshsfkwrwyhfdgtdwdqsrqlnkiyxf 180  
 QY 152 EVNRSNRNOEISGEYTTIATWTKDFPGRGNTSHSNFKRWYHFDGTDWDQSRQLNKIYKF 211

Db 181 rgtqkawdewdtengnydylymyanvdmhpevihelrnwgvvytntlnlqgfridavkh 240  
 QY 212 RGTQKAWDEWDIENGNYDYLYMAYADIMDHPEVINELRNWGVVYNTNLNLDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvnttggkpmfavaefwkndlgalenlntskwnhsvfdvplhynlyna 300  
 QY 272 IKYSYTRDOWLTHVNTTGGKPMFAVAEFWKNDLAAENLNTSKWNHSHVFDVPLHYNLYNA 331  
 Db 301 snsggyddmrnlngsvvqkphthavtfdvhdhdsqgealesfvqgwkplayalvltre 360  
 QY 332 SNSSGYDDMRNLNGSVVQKPHTHAVTFVDNHDSDQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidllqartqtfaygtqhdhdiigwtregness 420  
 QY 392 QGYPSVFGDYGIPTHGVPAMKSIDPLQARQTYAYGTQHDYFDHDIIGWTREGDSS 451  
 Db 421 hpnsglatimsdpggknkmyvgnkagqvrditgnrtgtvtinadgwnfsvnggsvs 480  
 QY 452 HPNSGLATIMSDPGGKNKMYVGNKAGQVVRDITGNRSQVTTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VVVKQ 516

## RESULT 14

ID W12124 standard; protein; 485 AA.  
 AC W12124;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D192N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 192 /label= D192N  
 FT WO9623873-Al.  
 PN 08-AUG-1996.  
 PD 05-FEB-1996; DK0056.  
 PF 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; l1lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12956, W12956, W12956, W12956, W12956,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R1835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hngtngtmqyfewylpndgnhwnrlrddaanlkskitavwippawktsqdndvgysa 60  
 QY 32 HNGTNGTMQYFEWYLPNDGNEHNRLLRDDAANLKSKITAVWIPPAAWKTSQNDVGYSa 91  
 Db 61 ydyldlqefnqktvrtkytrnqlgaavtslknngiqvgydvvmhkhgagdgteivnav 120  
 QY 92 YDYLDLQEFNQKTVRTKYTRNQLGAAVTSLKNNGIQVGYDVVMHKGAGDGTETEMWNAV 151  
 Db 121 evnrsnrngetsgeyaaleawtkfdpgrgnhshsfkwrwyhfdgtwdqsrqlgnkiyxf 180  
 QY 152 EVNRSNRNGETSGEYALEAWTKFDPGRGNTHSNFKRWYHFDGTDWDQSRQLGNKIYKF 211  
 Db 181 rgtgkawdewdtengnydylymyadvmhpevihelrnwgvvytntlnlqgfridavkh 240  
 QY 212 RGTGKAWDEWDIENGNYDYLYMAYADIMDHPEVINELRNWGVVYNTNLNLDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvnttggkpmfavaefwkndlgalenlntskwnhsvfdvplhynlyna 300  
 QY 272 IKYSYTRDOWLTHVNTTGGKPMFAVAEFWKNDLAAENLNTSKWNHSHVFDVPLHYNLYNA 331  
 Db 301 snsggyddmrnlngsvvqkphthavtfdvhdhdsqgealesfvqgwkplayalvltre 360  
 QY 332 SNSSGYDDMRNLNGSVVQKPHTHAVTFVDNHDSDQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidllqartqtfaygtqhdhdiigwtregness 420  
 QY 392 QGYPSVFGDYGIPTHGVPAMKSIDPLQARQTYAYGTQHDYFDHDIIGWTREGDSS 451  
 Db 421 hpnsglatimsdpggknkmyvgnkagqvrditgnrtgtvtinadgwnfsvnggsvs 480  
 QY 452 HPNSGLATIMSDPGGKNKMYVGNKAGQVVRDITGNRSQVTTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VVVKQ 516

## RESULT 15

ID W12108 standard; protein; 485 AA.  
 AC W12108;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant K269R.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 269 /label= K269R  
 FT WO9623873-Al.  
 PN 08-AUG-1996.  
 PD 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 10; ; l1lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12956, W12956, W12956, W12956, W12956,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R1835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic

CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match		91.3%;	Score 3536;	DB 20;	Length 485;
Best Local Similarity		94.8%;	Pred. No. 0.00e+00;		
Matches 460;		Conservative 19;	Mismatches 6;	Indels 0;	Gaps 0;
Db	1	hhngtnglmmyfeywlpndgnhnrlddaanlkskgitavwipawkgtsqndvgyga 60			
Qy	32	HHNGTNGTMMQYFEWHLPDNGNHNFLRDDAANLKSKITAVWIPAWKTSQNDVGYGA 91			
Db	61	ydlldlgefngktvrtkygttrnqlqaavtslknngiqvygdvvmnhkggadgteivnav 120			
Qy	92	YDLYDLGEFNGKGTVRTKYGTTRSQLQGAVTSLKNNGIQYGVVVMNHKGADGTEMVNAV 151			
Db	121	evnrnrnqetsgeyaieawtkfdipgrgnhssfkwrwyhfdgtwdqsrqlqnkiyxf 180			
Qy	152	EYVNRNRNQEISGEYIEAWTKFDFPGRGNTHSNFKRWYHFDGTDWDQSRQLQNKIYKF 211			
Db	181	rqtgkawdvevtdengnydlmivadmdhpevihelrnwgywytntlnldgfridavkh 240			
Qy	212	RGTGKAWDVEVDIENGNDYLMYADIDMDHPEVINELRNWGWYNTNLNDGFRIDAVKH 271			
Db	241	ikysfrdlwltvhrnttgkpmfavaefwrndlgalenylnktswnhsvfdvplhynlyna 300			
Qy	272	IKYSYTRDOWLTHVRNTTGKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYNLYNA 331			
Db	301	snsggydmrnlnilngsvqkphthavtfvnhdsqgealesfvqgwkfplayalvtre 360			
Qy	332	SNSGGYFDMRNILNGSVQKPHIHAVTFVDNHDSQPGEALESFVQSWFKPLAYALILTRE 391			
Db	361	ggypsvfgydygipthgypamkskldpllqarqtfaqtdhdyfdhhdliqwtregns 420			
Qy	392	QGYPVSFYGDYGIPTHGYPMSKSIDPLLQARQTYAGTQHDYFDHHDIIIGWTREGDSS 451			
Db	421	hpnsglatimsdpgpggnkwmvygknkagqvrdditgnrtgtvtinadgwnfsvngssvs 480			
Qy	452	HPNSGLATIMSDPGGNKWMYVGKKAGQVWRDITGNRSQTVTINADGWNFTVNGGAVS 511			
Db	481	vvvkq 485			
Qy	512	VWVKQ 516			

Search completed: Tue Sep 15 14:16:18 1998  
Job time : 150 secs.

\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:22:04 1998; MasPar time 7.52 Seconds

Tabular output not generated.  
484,408 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873

Sequence: 1 MKLHNRIISVLLTLLAVV.....ADGNGFTVNGAVSVWVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 77021 seqs, 7058996 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 33.562; Variance 161.785; scale 0.207

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2682	69.2	512	1	US-07-623- Sequence 5, Applicatio	6.31e-224
2	2680	69.2	512	1	US-08-720- Sequence 2, Applicatio	9.51e-224
3	2671	69.0	511	1	US-08-468- Sequence 35, Applicati	6.03e-223
4	2671	69.0	511	1	US-08-645- Sequence 3, Applicatio	6.03e-223
5	2671	69.0	511	2	PCT-US94-0 Sequence 33, Applicati	6.03e-223
6	2671	69.0	511	2	PCT-US95-1 Sequence 33, Applicati	6.03e-223
7	2662	68.7	512	1	US-07-623- Sequence 3, Applicatio	4.69e-222
8	2661	68.7	512	1	US-08-720- Sequence 6, Applicatio	1.80e-219
9	2632	68.0	483	2	PCT-US95-1 Sequence 32, Applicati	1.80e-219
10	2632	68.0	483	2	PCT-US94-0 Sequence 32, Applicati	1.80e-219
11	2632	68.0	483	1	US-08-468- Sequence 34, Applicati	1.80e-219
12	2632	68.0	487	2	PCT-US95-1 Sequence 37, Applicati	1.80e-219
13	2632	68.0	487	2	PCT-US94-0 Sequence 37, Applicati	1.80e-219
14	2629	67.9	483	1	US-08-645- Sequence 2, Applicatio	3.34e-219
15	2620	67.6	483	2	PCT-US95-1 Sequence 36, Applicati	2.11e-218
16	2620	67.6	483	2	PCT-US94-0 Sequence 36, Applicati	2.11e-218
17	2599	66.1	514	1	US-08-720- Sequence 4, Applicatio	5.75e-213
18	2599	66.1	520	1	US-08-645- Sequence 4, Applicatio	5.75e-213
19	2599	66.1	520	1	US-08-468- Sequence 36, Applicati	5.75e-213
20	2599	66.1	520	2	PCT-US95-1 Sequence 34, Applicati	5.75e-213
21	2599	66.1	520	2	PCT-US94-0 Sequence 34, Applicati	5.75e-213
22	2594	65.2	548	2	PCT-US94-0 Sequence 35, Applicati	7.54e-210
23	2524	65.2	548	1	US-08-645- Sequence 5, Applicatio	7.54e-210

24	2524	65.2	548	2	PCT-US95-1	Sequence 35, Applicati	7.54e-210
25	2524	65.2	548	1	US-08-468-	Sequence 37, Applicati	7.54e-210
26	1899	49.0	484	3	5171673-8	Patent No. 5171673.	2.83e-154
27	1899	49.0	487	3	5171673-6	Patent No. 5171673.	2.83e-154
28	284	7.3	60	3	5171673-4	Patent No. 5171673.	1.18e-13
29	174	4.5	38	3	5171673-2	Patent No. 5171673.	5.53e-05
30	138	3.6	428	1	US-08-343-	Sequence 4, Applicatio	2.55e-02
31	138	3.6	428	1	US-07-973-	Sequence 4, Applicatio	2.55e-02
32	138	3.6	434	1	US-08-184-	Sequence 6, Applicatio	2.55e-02
33	138	3.6	434	1	US-07-923-	Sequence 6, Applicatio	2.55e-02
34	135	3.5	687	1	US-08-470-	Sequence 10, Applicati	4.20e-02
35	135	3.5	687	1	US-08-467-	Sequence 10, Applicati	4.20e-02
36	132	3.4	687	1	US-08-204-	Sequence 10, Applicati	6.87e-02
37	127	3.3	478	1	US-08-720-	Sequence 7, Applicatio	1.56e-01
38	124	3.2	437	1	US-07-973-	Sequence 6, Applicatio	2.53e-01
39	124	3.2	437	1	US-08-343-	Sequence 6, Applicatio	2.53e-01
40	120	3.1	438	1	US-07-973-	Sequence 2, Applicatio	4.81e-01
41	120	3.1	438	1	US-08-343-	Sequence 2, Applicatio	4.81e-01
42	119	3.1	468	1	US-08-470-	Sequence 9, Applicatio	5.65e-01
43	119	3.1	468	1	US-08-204-	Sequence 8, Applicatio	5.65e-01
44	119	3.1	468	1	US-08-467-	Sequence 9, Applicatio	5.65e-01
45	114	2.9	468	1	US-08-467-	Sequence 6, Applicatio	1.25e+00

ALIGNMENTS

RESULT 1  
ID US-07-623-953-5 STANDARD; PRT; 512 AA.  
XX  
AC xxxxxx  
XX

Sequence 5, Application US/07623953

Sequence 5, Application US/07623953

Patent No. 5364782

GENERAL INFORMATION:

APPLICANT: Quax, Wilhelmus J

APPLICANT: Laroche, Yves

APPLICANT: Vollebregt, Adrianus W.H.

APPLICANT: Stanssens, Patrick

APPLICANT: Lauwereys, Marc

TITLE OF INVENTION: MUTANT MICROBIAL ALPHA-AMYLASES WITH

TITLE OF INVENTION: INCREASED THERMAL, ACID AND/OR ALKALINE STABILITY

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM

STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07/623,953

APPLICATION NUMBER: 19901129

FILING DATE: 19901129

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter, Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: GBRO-025/0005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-7622

TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 512 amino acids

Nov 15, 94

```

CC      TYPE: AMINO ACID
CC      TOPOLOGY: linear
SQ      SEQUENCE 512 AA; 58492 MW; 1370525 CN;

Query Match
Best Local Similarity 69.2%; Score 2682; DB 1; Length 512;
Matches 347; Conservative 86; Mismatches 76; Indels 10; Gaps 7;

Db 1  MKQQRKRYARLTLTLLFALIFLLPH-SAAAAA--N-LNGTLMQYFEWYMPNDGQHWKRLQN 56
QY 1  MKLHNRIISVLLTLLAVAVLPFYMPTEPAQAHHNGTNGTMMQYFEHLLPNDGNHNRRLD 60

Db 57  DSAYLAEGITAVTTPPAYKGTSDQADVGAYDLDLGEFHQKGTVRTKYTGKELQSAI 116
QY 61  DAANKSKGIIAVTTPPAYKGTSDQADVGAYDLDLGEFNOKGTVRTKYTRSLOQAV 120

Db 117  KSLHSRDINVTGVVNHKGGADATEDVTAVEVDPADNRNVISGEHLIKAWTHFFHPGRG 176
QY 121  TSLKNGIQVYGDVVMNHKGGADGTENVNVEVNSNRNQEISGEYTTIAWTKFDFPGRG 180

Db 177  STYSDFKWHVTHFDGTDWDESKL-NRIYKFG--KAWDWEVSNENGVNHYREKTKGEMFTVAEYQ 233
QY 181  NTHSNFKRWYHFDGTDWDSRQLQNKIKYKFRGTCKAWDEVDIENGVDYLYADIDMD 240

Db 234  HPDVAAEIKRWGTWYANELQDGFRLDAVKHKFSLRDWNVNHYREKTKGEMFTVAEYQ 293
QY 241  HPEVINELRNKGWVYITNLNDGFRIDAVKHKISYTRDMLTHVRNTTGKPMFAVEFWK 300

Db 294  NDGALENYLNKTNFNSHVSFVPLHYQFHAASQTGGGYDMRKLNGTVVSKHPLKSVTFV 353
QY 301  NDAAIENYLNKTSWNHVSFVPLHYNLNASNGGYFDMRNILNGSVVQKPHAVTFV 360

Db 354  DNHDTPQOSLESTVQTFKFLAYAFILTRSGYPOVYFDMYGTGKDSQREIPALKHKI 413
QY 361  DNHDSPQGEALESFVQSWFKFLAYALILTRQGYPSVFGYDYG--PTH-GVPSMKS 417

Db 414  EPIKARKQYAGAHQYDFDHDIVGTREGDSSVANSGLAALITDGGPGAKRYVYGRN 473
QY 418  DPLQARQTYAGTQHDYFDHDIIGTREGDSSHPNSGLATIMSDGPGGNKMYVGKHK 477

Db 474  AGETWHDITGNRSEPVVINSSEGVHNGSVSIYVQR 512
QY 478  AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVVVKQ 516

RESULT 2
ID  US-08-720-899-2      STANDARD;      PRT;      512 AA.
XX  xxxxxx
AC  xxxxxx
XX  xxxxxx

Sequence 2, Application US/08720899

Sequence 2, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
CC  APPLICANT: Bisgaard-Frantzen, Henrik
CC  APPLICANT: Borchert, Torben Vedel
CC  APPLICANT: Svendsen, Allan
CC  APPLICANT: Thellersen, Marianne
CC  APPLICANT: Van der Zee, Pia
CC  TITLE OF INVENTION: AMYLASE VARIANTS
CC  NUMBER OF SEQUENCES: 38
CC  CORRESPONDENCE ADDRESS:
CC  ADDRESSEE: No. 5753460o No. 5753460disk of No. 5753460th America, Inc.
CC  STREET: 405 Lexington Avenue, 64th Floor
CC  CITY: New York
CC  STATE: New York
CC  COUNTRY: USA
CC  ZIP: 10174-6401
CC  COMPUTER READABLE FORM:
CC  MEDIUM TYPE: Floppy disk

```

```

CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/720,899
CC      FILING DATE: 10-OCT-1996
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/343,804
CC      FILING DATE: 22-NOV-1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Lowney Dr., Karen A.
CC      REGISTRATION NUMBER: 31,274
CC      REFERENCE/DOCKET NUMBER: 4054,214-US
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 212-867-0123
CC      TELEFAX: 212-878-9655
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 512 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 512 AA; 58520 MW; 1370005 CN;

Query Match
Best Local Similarity 69.2%; Score 2680; DB 1; Length 512;
Matches 347; Conservative 85; Mismatches 77; Indels 10; Gaps 7;

Db 1  MKQQRKRYARLTLTLLFALIFLLPH-SAAAAA--N-LNGTLMQYFEWYMPNDGQHWKRLQN 56
QY 1  MKLHNRIISVLLTLLAVAVLPFYMPTEPAQAHHNGTNGTMMQYFEHLLPNDGNHNRRLD 60

Db 57  DSAYLAEGITAVTTPPAYKGTSDQADVGAYDLDLGEFHQKGTVRTKYTGKELQSAI 116
QY 61  DAANKSKGIIAVTTPPAYKGTSDQADVGAYDLDLGEFNOKGTVRTKYTRSLOQAV 120

Db 117  KSLHSRDINVTGVVNHKGGADATEDVTAVEVDPADNRNVISGEHLIKAWTHFFHPGRG 176
QY 121  TSLKNGIQVYGDVVMNHKGGADGTENVNVEVNSNRNQEISGEYTTIAWTKFDFPGRG 180

Db 177  STYSDFKWHVTHFDGTDWDESKL-NRIYKFG--KAWDWEVSNENGVNHYREKTKGEMFTVAEYQ 233
QY 181  NTHSNFKRWYHFDGTDWDSRQLQNKIKYKFRGTCKAWDEVDIENGVDYLYADIDMD 240

Db 234  HPDVAAEIKRWGTWYANELQDGFRLDAVKHKFSLRDWNVNHYREKTKGEMFTVAEYQ 293
QY 241  HPEVINELRNKGWVYITNLNDGFRIDAVKHKISYTRDMLTHVRNTTGKPMFAVEFWK 300

Db 294  NDGALENYLNKTNFNSHVSFVPLHYQFHAASQTGGGYDMRKLNGTVVSKHPLKSVTFV 353
QY 301  NDAAIENYLNKTSWNHVSFVPLHYNLNASNGGYFDMRNILNGSVVQKPHAVTFV 360

Db 354  DNHDTPQOSLESTVQTFKFLAYAFILTRSGYPOVYFDMYGTGKDSQREIPALKHKI 413
QY 361  DNHDSPQGEALESFVQSWFKFLAYALILTRQGYPSVFGYDYG--PTH-GVPSMKS 417

Db 414  EPIKARKQYAGAHQYDFDHDIVGTREGDSSVANSGLAALITDGGPGAKRYVYGRN 473
QY 418  DPLQARQTYAGTQHDYFDHDIIGTREGDSSHPNSGLATIMSDGPGGNKMYVGKHK 477

Db 474  AGETWHDITGNRSEPVVINSSEGVHNGSVSIYVQR 512
QY 478  AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVVVKQ 516

RESULT 3
ID  US-08-468-700-35      STANDARD;      PRT;      511 AA.
XX  xxxxxx
AC  xxxxxx
XX  xxxxxx
DT  xxxxxx
XX  xxxxxx

```



QY 361 DNHDSQPGEALSFVQSWFKPLAYALILITREGQYPSVFGDYGI--PTH-GVPSMKSKI 417  
 Db 413 EPIKARKQYAGAOHDYFDHDIIVGWTREGDSSVANSGLAALITDGPAGAKRMVVGRO 472  
 QY 418 DPLQARQYAYGTQHDYFDHDIIVGWTREGDSSHPNSGLATIMSDGPGGNKMWVGRHK 477  
 Db 473 AGETHDITGNRSEPVVINSSEGEFHVNGGVSIIYVOR 511  
 QY 478 AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVWVKQ 516

RESULT 5  
 ID PCT-US94-01553A-33 STANDARD; PRT; 511 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 33, Application PC/TUS9401553A  
 Sequence 33, Application PC/TUS9401553A  
 GENERAL INFORMATION:  
 APPLICANT: GENENCOR INTERNATIONAL, INC.  
 TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
 NUMBER OF SEQUENCES: 68  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International, Inc.  
 STREET: 180 Kimball Way  
 CITY: South San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/01553A  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Horn, Margaret A.  
 REGISTRATION NUMBER: 33,401  
 REFERENCE/DOCKET NUMBER: GC220-2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 742-7536  
 TELEFAX: (415) 742-7217  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 511 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE 511 AA; 58364 MW; 1365410 CN;

Query Match 69.0%; Score 2671; DB 2; Length 511;  
 Best Local Similarity 66.9%; Pred. No. 6.03e-223;  
 Matches 347; Conservative 86; Mismatches 75; Indels 11; Gaps 8;

Db 1 MKQQRRLVRLTLFLALIFLPH-SAAAA--N-INGTLMQYFEWMPNDG-HWKRLQN 55  
 QY 1 MKLHNRRIISVLLTLAVLFPYMTPEAQAHNGTNGTMMQYFEWHLNDGNHNRRLD 60  
 Db 56 DSAYLAEGITAVTWPAPKGTQADYGVAYDLDLGEFHQKGVRYKYGKELQSAI 115  
 QY 61 DAANLKSIGITAVTWPAPKGTQADYGVAYDLDLGEFNQKGVRYKYGTRSOLOGAV 120  
 Db 116 KSLSRDINVDVGVVNHKGGADATEDTAVEVDPADNRNVISGEHLKAWTHFHPGRG 175  
 QY 121 TSLKNGNQVYGDVVMNHKGGADGTEMVNAVENVNRRNQELSGEYTTIETAWTKFDFPGRG 180

Db 176 STYDFKWHWFHFDGTDWDESKL-NRIYKFOG--KAWDWEVSNENGNNDYDLMYADIDYD 232  
 QY 181 NTHSNFKRWTHFDGTDWDSQSLQNKIYKFRGTGKAWDWEVDIENGNDYDLMYADIDMD 240  
 Db 233 HPDVAAEIKRMTWYANELQLDGFRLDAVKHIKFSFLRDVWVHVREKTKGEMFTVAEYHQ 292  
 QY 241 HPEVINELRNWGVWYNTNLNDGFRIDAVKHIKYSYTRDLWLTHTVRNTTCKPMPFAVEFWK 300  
 Db 293 NDLAGALENLYLNKTNNHNSVDFVPLHYQFHAASSTOGGDMRKLNGTVVSKHPLKSVTFV 352  
 QY 301 NDLAANIENLYLNKTSWNHNSVDFVPLHYNLYNASNGGYEDMRNLNGSVVQKHPHIAVTFV 360  
 Db 353 DNHDTPQGSLESTVOTWFKPLAYAFILITRESGYPOVYGYDMYGTGDSQRIKPKHKI 412  
 QY 361 DNHDSQPGEALSFVQSWFKPLAYALILITREGQYPSVFGDYGI--PTH-GVPSMKSKI 417  
 Db 413 EPIKARKQYAGAOHDYFDHDIIVGWTREGDSSVANSGLAALITDGPAGAKRMVVGRO 472  
 QY 418 DPLQARQYAYGTQHDYFDHDIIVGWTREGDSSHPNSGLATIMSDGPGGNKMWVGRHK 477  
 Db 473 AGETHDITGNRSEPVVINSSEGEFHVNGGVSIIYVOR 511  
 QY 478 AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVWVKQ 516

RESULT 6  
 ID PCT-US95-10426-33 STANDARD; PRT; 511 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 33, Application PC/TUS9510426  
 Sequence 33, Application PC/TUS9510426  
 GENERAL INFORMATION:  
 APPLICANT: GENENCOR INTERNATIONAL, INC.  
 TITLE OF INVENTION: An Improved Cleaning Composition  
 NUMBER OF SEQUENCES: 68  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International  
 STREET: 180 Kimball Way  
 CITY: South San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/10426  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stone, Christopher L.  
 REGISTRATION NUMBER: 33,401  
 REFERENCE/DOCKET NUMBER: GC220-3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 742-7536  
 TELEFAX: (415) 742-7217  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 511 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE 511 AA; 58364 MW; 1365410 CN;

Query Match 69.0%; Score 2671; DB 2; Length 511;  
 Best Local Similarity 66.9%; Pred. No. 6.03e-223;



[illegible]

CC	FILING DATE:	19901129
CC	CLASSIFICATION:	435
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Rae-Venter, Barbara
CC	REGISTRATION NUMBER:	32,750
CC	REFERENCE/DOCKET NUMBER:	GBRO-025/00US
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	415-494-7622
CC	TELEFAX:	415-857-0663
CC	TELEX:	380816 COOLEY PA
CC	INFORMATION FOR SEQ ID NO:	3:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	512 amino acids
CC	TYPE:	AMINO ACID
CC	TOPOLOGY:	linear
SQ	SEQUENCE	512 AA; 58452 MW; 1370688 CN;
Query Match            68.7%; Score 2662; DB 1; Length 512;		
Best Local Similarity 66.5%; Pred. No. 3.82e+22;		
Matches     345; Conservative     87; Mismatches 77; Indels 10; Gaps		
Db	1	MKQOKRLYARSVTLFLFALILPH-SAAAAA--N-LNCTLMQYFEWYMPNDGQHWRQLN 56
Qy	1	MKLHNRISVLLTLLAVAVLFPMTEPAQAHHNGINGNTMMQFEWHLPLPDGNHWRLRD 60
Db	57	DSAYLAERGITAVWIPPAYXGTSGQADVGYGAYLDLGEFHQXGTVRTKYGTGELQSAI 116
Qy	61	DAANLKSGKITAWWPAPWKAGTSQNDVGYGAYLDLGEFNQKGTVRTKYGTRSLOQGV 120
Db	117	KSLHSRDINYGDVVINHKGADATEDVTAVEVDPADNRNVISGEHLIKAWTHFFPGRG 176
Qy	121	TSLNKNGIQVYGDVMNHKGADGTEMNAVEVNRSNRNQEISGEVTIEAWTKDFDPGRG 180
Db	177	STYSDFWHYHFDGTDWDERSKL-NRIYFKQG--RAWDEWVENGENGYDLYMADIDYD 233
Qy	181	NTHSNFWRWYHFDGTDWDQSRQLNKIYKFRTGKAWDWEVDIENGNDYLYMADIMD 240
Db	234	HPDVAAEKRWGTWAYNELQLDGFRDLDAVRHKIFSFLRDMWNHVREKTGEMFTVAEYWQ 293
Qy	241	HPEVINELRWGWYYTNLTNLGFRIDAVERHIKYSYTRDWLTHVRNTTGKPMFAVAEFWK 300
Db	294	NDLGALENYLKNFNHNSVFDPVLHYQFHAASTQGGGYDMRKLLNGTIVSKHPLKSVTFV 353
Qy	301	NDLAAENLYNLKTSWNHSFVDPVLHYNLYNASNGGYFDMRNILINGSVVOKHPHAATFV 360
Db	354	DNHDTQPQSILESTVQWFPLAYAFILTTRESGPQVFGMDYGTGKDSORETPALKHKI 413
Qy	361	DNHDSQPGEALESVSQWFKPLAYALLTREQGPSVFGDYGI--PTH-GYPMSKSKI 417
Db	414	EPILKARKQYAGAQHDYFOHHDIWGVTREGSDSVANSGLAALITDGPGGAKRMVYGRON 473
Qy	418	DPLLQARQTAYGTQHDYFDDHDIIGWTRGDSHPNSGLATMSDGPGGKKMYVGGKH 477
Db	474	AGETVHDDITGNRPVPVINSEGMEGFHVNGGSVIYVQR 512
Qy	478	AGQVWRDITGNRSCTVTINADGMGNETVNGGAVSWVKQ 516
RESULT	8	
ID	US-08-720-899-6	STANDARD; PRT; 549 AA.
XX	xxxxxx	
AC		
CC		
DT		
XX		
DE	Sequence 6, Application US/08720899	
CC	Sequence 6, Application US/08720899	
CC	Patent No. 5753460	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Bisgaard-Frantzen, Henrik	
CC	APPLICANT: Borchert, Torben Vedel	
CC	APPLICANT: Svendsen, Allan	

CC APPLICANT: Thellersen, Marianne  
CC APPLICANT: Van der Zee, Pia  
CC TITLE OF INVENTION: AMYLASE VARIANTS  
CC NUMBER OF SEQUENCES: 38  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.  
CC STREET: 405 Lexington Avenue, 64th Floor  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10174-6401  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/720,899  
CC FILING DATE: 10-OCT-1996  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/343,804  
CC FILING DATE: 22-NOV-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Lowney Dr. Karen A.  
CC REGISTRATION NUMBER: 31,274  
CC REFERENCE/DOCKET NUMBER: 4054, 214-US  
CC TELEPHONE: 212-867-0123  
CC TELEFAX: 212-878-9655  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 549 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 549 AA; 62566 MW; 1687311 CN;

Query Match 68.7%; Score 2661; DB 1; Length 549;  
Best Local Similarity 67.5%; Pred. No. 4.69e-222;  
Matches 341; Conservative 78; Mismatches 81; Indels 5; Gaps 5;  
DB 15 LLAPLLVTS-LFCPTGPAPAA-APEFNGTMMOYFEWLPDDGTMTKVAENLSSLSGI 72  
QY 11 LLLTLLAVAVLPYMTPEPAQHNGTNGTMMOYFEWHLPGDNGNHNRLRDAANLKSNGI 70  
DB 73 TALWLPAYKGTSSRDYGYDLYDLGEFNGKGTGRTKTKAQYQLQ-AIQAAHAAGMQ 131  
QY 71 TAVWIPPAWKGTSONDVGAYDLYDLGEFNGKGTGRTKTKGTRSQ-LOGAVTSLKNGIQ 129  
DB 132 VYADVDFHKGAGDGTWVDVAVEVNPDRNOEISGTQIOAWTKFDPGRGNTYSFKWR 191  
QY 130 YGDDVNNHKGAGDGTWVNAVEVNSRNOEISGEYTIKFTGRTGRTGRTGRTGRTGRT 189  
DB 192 WYHFDGVDWDESRKL-SRIYKFRIGKAWDWEVDTEGNDYDLYMYADLMDHPEVTELK 250  
QY 190 WYHFDGTDQSQRLQNKYIKFRTGKAWDWEVDTEGNDYDLYMYADLMDHPEVTELK 249  
DB 251 NWGKVVNTNIDGFRDLDAVKKHFKSFPDMLSVRSQTKPLFTVGEYWSYDINKLHNY 310  
QY 250 NWGKVVNTNIDGFRDLDAVKKHFKSFPDMLSVRSQTKPLFTVGEYWSYDINKLHNY 309  
DB 311 ITKDTGMSLFDAPLHKKETASKSGAGFDMRLMTNLMKDDPTLATVFDVNDHDEPGQ 370  
QY 310 LKNTSNHNSVDFVPLHNLNNAHNSGGYFDMRLNLSVQKHPHIAVTFVNDHDSQGE 369  
DB 371 ALOSVDPMFKPLAYAFILTRQEGYPCVFGYDYGIPQYINPSLKSIDKIDPLLIARDYAY 430  
QY 370 ALESFVQSKFKPLAYALITREQYVSVFYDYGIPTHGVPSMKSIDKIDPLLIARDYAY 429  
DB 431 GTQHDYLDHSDIIGWTREGTEKPGSGLAALITDGGGSKWMTYVQKHAKGVYDITGNR 490  
QY 430 GTQHDYLDHSDIIGWTREGSSHPNSGLATIMSDGPGGNKMYVYGHKAGQVYDITGNR 489

DB 491 SDVTITNSDGGGEFFKVGSGSVSWV 515  
QY 490 SGTVTINADGNGFTVNGGAVSWV 514  
RESULT 9  
ID PCT-US95-10426-32 STANDARD; PRT; 483 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
XX  
Sequence 32, Application PC/TUS9510426  
XX  
XX  
Sequence 32, Application PC/TUS9510426  
CC GENERAL INFORMATION:  
CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
CC TITLE OF INVENTION: An Improved Cleaning Composition  
CC NUMBER OF SEQUENCES: 68  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genencor International  
CC STREET: 180 Kimball Way  
CC CITY: South San Francisco  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/10426  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: STONE, Christopher L.  
CC REGISTRATION NUMBER: 33,401  
CC REFERENCE/DOCKET NUMBER: GC220-3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 742-7536  
CC TELEFAX: (415) 742-7217  
CC INFORMATION FOR SEQ ID NO: 32:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 483 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 483 AA; 55211 MW; 1221290 CN;  
Query Match 68.0%; Score 2632; DB 2; Length 483;  
Best Local Similarity 68.9%; Pred. No. 1.80e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;  
DB 4 NGTLMQYFWMYMPNDGQHWKRLONDSAYLABGHITAVWIPPAYKGTQADYGYGAYDLYD 63  
QY 37 NGTMMQYFWEHLPGDNGNHNRLRDAANLKSNGITAVWIPPAYKGTQADYGYGAYDLYD 96  
DB 64 LGFHFQKGTGRTKTKGELQSAIKSLHSRDIYVGVVNNHKGAGDTEVTAIVEDPA 123  
QY 97 LGFHFQKGTGRTKTKGELQSAIKSLHSRDIYVGVVNNHKGAGDTEVTAIVEDPA 156  
DB 124 DRNRVTSGEHLIKAWTHFHPGRGSTYSDPKWHYHFDGTDWDESRKL-NRIYKFG--K 180  
QY 157 NRQETSGEYTIKFTGRTGRTGRTGRTGRTGRTGRTGRTGRTGRTGRTGRTGRTGRT 216  
DB 181 AWDWEVSNNGNDYDLYMYADIDYDHPDVAABEIKRWGTYWYANELQDGLDGLDAVKHIFSF 240  
QY 217 AWDWEVDIENGNDYDLYMYADIDMDHPEVINELRWGTYWYANELQDGLDGLDAVKHIFSF 276  
DB 241 LRDMVNHVREKTKMETVAVYQNDLGALENLKNKTNFHNHSDVEDPLHYQFHAASQTGG 300

QY	277	TRDLWLRHVRNTTQKPF	FAVFAE	WKNDLAA	IENTLNKTSWNH	SFVDFPLHYNL	YNASNSG	336					
Db	301	GYDMRKLLNGTVYSK	PKLSP	FVNDH	DTQPGOS	LESTVQTWFK	PLAYAFIL	TREGGYQ 360					
QY	337	YFDMRNILNGSVVQK	PIH	AVFVND	HDSQPG	EALSFVQSWFK	PLAYALL	TREGGYPS 396					
Db	361	VFYGDMYGTGDSQ	REIPALK	KHIEP	LKARKO	YAGAH	DFDHDH	IYGVWTR	REGDSSVA 420				
QY	397	VFYGDYGI--PTH	-GVPS	MKS	KIDPL	LQARQ	TYAYGTQ	HDYFD	HDHDIIGWTR	REGDSSHP 453			
Db	421	NSGLAALITD	PGGAKRM	YVGR	ONAGET	WHDIT	GNSE	PVWIN	SEGS	GFHVNGGS	VSYI 480		
QY	454	NSGLATIMSD	PGGKNW	MYGK	HKHGA	GOVWR	DIITGN	RSRGT	VTINAD	GWGNFT	VNGGAVSW 513		
Db	481	VQR	483										
QY	514	VQK	516										
RESULT	10												
ID	PCT-US94-01553A-32	STANDARD;	PRT;	483	AA.								
XX	xxxxxx												
XX													
DT													
XX													
DE													
CC	Sequence 32, Application PC/TUS9401553A												
CC	Sequence 32, Application PC/TUS9401553A												
CC	GENERAL INFORMATION:												
CC	APPLICANT: GENENCOR INTERNATIONAL, INC.												
CC	TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase												
CC	NUMBER OF SEQUENCES: 68												
CC	CORRESPONDENCE ADDRESS:												
CC	ADDRESSEE: Genencor International, Inc.												
CC	STREET: 180 Kimball Way												
CC	CITY: South San Francisco												
CC	STATE: CA												
CC	COUNTRY: USA												
CC	ZIP: 94080												
CC	COMPUTER READABLE FORM:												
CC	MEDIUM TYPE: Floppy disk												
CC	COMPUTER: IBM PC compatible												
CC	OPERATING SYSTEM: PC-DOS/MS-DOS												
CC	SOFTWARE: Patent In Release #1.0, Version #1.25												
CC	CURRENT APPLICATION DATA:												
CC	APPLICATION NUMBER: PCT/US94/01553A												
CC	FILING DATE:												
CC	CLASSIFICATION:												
CC	ATTORNEY/AGENT INFORMATION:												
CC	NAME: Horn, Margaret A.												
CC	REGISTRATION NUMBER: 33,401												
CC	REFERENCE/DOCKET NUMBER: GC220-2												
CC	TELECOMMUNICATION INFORMATION:												
CC	TELEPHONE: (415) 742-7536												
CC	TELEFAX: (415) 742-7217												
CC	INFORMATION FOR SEQ ID NO: 32:												
CC	SEQUENCE CHARACTERISTICS:												
CC	LENGTH: 483 amino acids												
CC	TYPE: amino acid												
CC	STRANDEDNESS: single												
CC	TOPOLOGY: linear												
CC	MOLECULE TYPE: protein												
CC	SEQUENCE 483 AA; 55211 MW; 1221290 CN;												
Query Match 68.0%; Score 2632; DB 2; Length 483;													
Best Local Similarity 68.9%; Pred. No. 1.80e-219;													
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps													
Db	4	NGTLMQYFEW	MPNDG	QHW	RLQ	ND	SA	YLAE	HGIT	AVWIP	PAYKGT	SQADVGY	GAYDLXD 63
QY	3												

```
CC TELEPHONE: (415) 742-7555
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 34:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 483 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC SEQUENCE 483 AA; 55211 MW; 1221290 CN;

Query Match 68.0%; Score 2632; DB 1; Length 483;
Best Local Similarity 68.9%; Pred. No. 1.80e-219;
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;

Db 4 NGTLMQYFEWMPNDGQHWKRLQNDLSAYLAEHGITAIVWIPPAKGTQADYGYDAYDLYD 63
QY 37 NGTMMQYFEWHLNDGNHNRRLRDDAANLKSIGITAVWIPPAKGTQSDYGYDAYDLYD 96
Db 64 LGFPHQKGTVRTKYGTGKELQSAIKSLHSRDIINVGDVNVNHNKGGADATEDVTAVEVDPA 123
QY 97 LGFENKGTVRTKYGTGKELQSAIKSLHSRDIINVGDVNVNHNKGGADATEDVTAVEVDPA 156
Db 124 DNRNVSIGHEHLIKAWTHFHPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFGQ--K 180
QY 157 NRNQEISGEYTIKAWTKDFPGRGNTHSNFKRWYHFDGTDWDSRKLQNKIYKFRGTGK 216
Db 181 AWDWEVSNENGNIDYLYADIDYDHPDVAAEIKRWGTWYANELQDLQDFRLDAVKHKIYKFSF 240
QY 217 AWDWEVDIENGNDYLYADIDYDHPDVAAEIKRWGTWYANELQDLQDFRLDAVKHKIYKFSF 276
Db 241 LRDVNVNHNKGGADATEDVTAVEVDPA 300
QY 277 TRDMLTHVRNTTGPMPFAVEFKNLAAIENYLNKTWNHSHVDFVPLHYNLNASNGG 336
Db 301 GYDMRLLNGTVSKHPLKSVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGSYPQ 360
QY 337 YFDMRNLNGSVVQKHPHIAVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGSYPQ 396
Db 361 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYCTQHDYDFDHHDIIGWTRGDSHP 420
QY 397 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYCTQHDYDFDHHDIIGWTRGDSHP 453
Db 421 NSGLAALITDGGPGAKRMVYGRONAGTWHDTGNRSEPVVINSSEGWGEFHVNGGSVSIY 480
QY 454 NSGLATIMSDGPGGNKMWYGVKHKAGQVWRDITGNRSGTGTINADGWNFTVNGGAVSVW 513
Db 481 VQR 483
QY 514 VKQ 516

RESULT 12
ID PCT-US95-10426-37 STANDARD; PRT; 487 AA.
XX
AC xxxxxx
XX

Sequence 37, Application PC/TUS9510426
Sequence 37, Application PC/TUS9510426
GENERAL INFORMATION:
CC APPLICANT: GENENCOR INTERNATIONAL, INC.
CC TITLE OF INVENTION: An Improved Cleaning Composition
CC NUMBER OF SEQUENCES: 68
CC CORESPONDENCE ADDRESS:
CC ADDRESSEE: Genencor International
CC STREET: 180 Kimball Way
CC CITY: South San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94080
```

```
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10426
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STONE, Christopher L.
CC REGISTRATION NUMBER: 33,401
CC REFERENCE/DOCKET NUMBER: GC220-3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 742-7536
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 37:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 487 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 487 AA; 55495 MW; 1241380 CN;

Query Match 68.0%; Score 2632; DB 2; Length 487;
Best Local Similarity 68.9%; Pred. No. 1.80e-219;
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;

Db 8 NGTLMQYFEWMPNDGQHWKRLQNDLSAYLAEHGITAIVWIPPAKGTQADYGYDAYDLYD 67
QY 37 NGTMMQYFEWHLNDGNHNRRLRDDAANLKSIGITAVWIPPAKGTQSDYGYDAYDLYD 96
Db 68 LGFPHQKGTVRTKYGTGKELQSAIKSLHSRDIINVGDVNVNHNKGGADATEDVTAVEVDPA 127
QY 97 LGFENKGTVRTKYGTGKELQSAIKSLHSRDIINVGDVNVNHNKGGADATEDVTAVEVDPA 156
Db 128 DNRNVSIGHEHLIKAWTHFHPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFGQ--K 184
QY 157 NRNQEISGEYTIKAWTKDFPGRGNTHSNFKRWYHFDGTDWDSRKLQNKIYKFRGTGK 216
Db 185 AWDWEVSNENGNIDYLYADIDYDHPDVAAEIKRWGTWYANELQDLQDFRLDAVKHKIYKFSF 244
QY 217 AWDWEVDIENGNDYLYADIDYDHPDVAAEIKRWGTWYANELQDLQDFRLDAVKHKIYKFSF 276
Db 245 LRDVNVNHNKGGADATEDVTAVEVDPA 304
QY 277 TRDMLTHVRNTTGPMPFAVEFKNLAAIENYLNKTWNHSHVDFVPLHYNLNASNGG 336
Db 305 GYDMRLLNGTVSKHPLKSVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGSYPQ 364
QY 337 YFDMRNLNGSVVQKHPHIAVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGSYPQ 396
Db 365 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYCTQHDYDFDHHDIIGWTRGDSHP 424
QY 397 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYCTQHDYDFDHHDIIGWTRGDSHP 453
Db 425 NSGLAALITDGGPGAKRMVYGRONAGTWHDTGNRSEPVVINSSEGWGEFHVNGGSVSIY 484
QY 454 NSGLATIMSDGPGGNKMWYGVKHKAGQVWRDITGNRSGTGTINADGWNFTVNGGAVSVW 513
Db 485 VQR 487
QY 514 VKQ 516

RESULT 13
ID PCT-US94-01553A-37 STANDARD; PRT; 487 AA.
XX
AC xxxxxx
XX
DT
XX
```

DE Sequence 37, Application PC/TUS9401553A  
XX Sequence 37, Application PC/TUS9401553A  
CC GENERAL INFORMATION:  
CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
CC TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
CC NUMBER OF SEQUENCES: 68  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genencor International, Inc.  
CC STREET: 180 Kimball Way  
CC CITY: South San Francisco  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US94/01553A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Horn, Margaret A.  
CC REGISTRATION NUMBER: 33,401  
CC REFERENCE/DOCKET NUMBER: GC220-2  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 742-7536  
CC TELEFAX: (415) 742-7217  
CC INFORMATION FOR SEQ ID NO: 37:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 487 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC ORGANISM: Bacillus licheniformis  
CC SEQUENCE 487 AA; 55495 MW; 1241380 CN;

Query Match 68.0%; Score 2632; DB 2; Length 487;  
Best Local Similarity 68.9%; Pred. No. 1,80e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;  
Db 8 NGLTMOYFEWYMPNDGQHWKRLQND SAYLA EHGITAVIPPA YKGT SQADVGYGAYDLYD 67  
QY 37 NGTMMQYFEWHLPNDCGNHWNRLRDDAANLKSXGITAVIPPA YKGT SQNDVGYGAYDLYD 96  
Db 68 LGEFHQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTA VEVDPA 127  
QY 97 LGEFNQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTA VEVDPA 156  
Db 128 DNRNVISGEHLKAWTHFFPGRGSTYSDFKWHYHFDGTDWDESRKL-NRIYKFGQ--K 184  
QY 157 NRNQEISGEYTI EATWKTFDFPGRGNTSHFNKRWYHFDGTDWQSQRLQNKIYKFRGTGK 216  
Db 185 AWDWEYSNENGYDLYMADIDYDHPDVA AEIKRWGTWYANELQLDGFRLDAVKHKFSF 244  
QY 217 AWDWEVDIENGVDYLMYADIDMDHPEVINELRNNGVWYTNLTNLGDFRIDAVKHKIYSY 276  
Db 245 LRDVNVHREKTKGEMFTVAEYVNDLGALENYLNKTNFNSHVSFVPLHYQFHAASTQGG 304  
QY 277 TRDWLTHVRNTTGKPMFAVAEFWKNDLAAIENLNKTSWNHSHVSFVPLHYLNASNSGG 336  
Db 305 GYDMRKLNGTVSKHPLKAVTFVDNHDTPQGSLSTVQTWFKPLAYAFILTRREGSYPQ 364  
QY 337 YFDMRNILGVSQVQKHPITHAVTFVDNHDTPQSGALSFVQSWFKPLAYALILTRREGYPS 396  
Db 365 VFYDGMGTGKDSQREIPALKHKIEPILKARKQYAYCAQHDYFDHHDIVGWTREGDSSVA 424  
QY 397 VFYDGYGI--PTH-GVPSMKSIDPLQARQTYAGTQHDYFDHHDIIIGWTREGDSSHP 453  
Db 425 NSGLATIMSDPGCGNKWYVCKHKGAGQVVRDITGNRSCTVTINADGWNFTVNGGAVSVW 513

QY 454 NSGLATIMSDPGCGNKWYVCKHKGAGQVVRDITGNRSCTVTINADGWNFTVNGGAVSVW 513  
Db 485 VQR 487  
QY 514 VKQ 516  
RESULT 14  
ID US-08-645-971-2 STANDARD; PRT; 483 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 2, Application US/08645971  
XX  
CC Sequence 2, Application US/08645971  
CC Patent No. 5763385  
CC GENERAL INFORMATION:  
CC APPLICANT:  
CC TITLE OF INVENTION: Modified Alpha-Amylases Having Altered  
CC TITLE OF INVENTION: Calcium Binding Properties  
CC NUMBER OF SEQUENCES: 5  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/645,971  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 483 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC ORIGINAL SOURCE:  
CC ORGANISM: Bacillus licheniformis  
CC SEQUENCE 483 AA; 55268 MW; 1217764 CN;

Query Match 67.9%; Score 2629; DB 1; Length 483;  
Best Local Similarity 68.9%; Pred. No. 3,34e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;  
Db 4 NGLTMOYFEWYMPNDGQHWKRLQND SAYLA EHGITAVIPPA YKGT SQADVGYGAYDLYD 63  
QY 37 NGTMMQYFEWHLPNDCGNHWNRLRDDAANLKSXGITAVIPPA YKGT SQNDVGYGAYDLYD 96  
Db 64 LGEFHQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTA VEVDPA 123  
QY 97 LGEFNQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTA VEVDPA 156  
Db 124 DNRNVISGEHLKAWTHFFPGRGSTYSDFKWHYHFDGTDWDESRKL-NRIYKFGQ--K 180  
QY 157 NRNQEISGEYTI EATWKTFDFPGRGNTSHFNKRWYHFDGTDWQSQRLQNKIYKFRGTGK 216  
Db 181 AWDWEYSNENGYDLYMADIDYDHPDVA AEIKRWGTWYANELQLDGFRLDAVKHKFSF 240  
QY 217 AWDWEVDIENGVDYLMYADIDMDHPEVINELRNNGVWYTNLTNLGDFRIDAVKHKIYSY 276  
Db 241 LRDVNVHREKTKGEMFTVAEYVNDLGALENYLNKTNFNSHVSFVPLHYQFHAASTQGG 300  
QY 277 TRDWLTHVRNTTGKPMFAVAEFWKNDLAAIENLNKTSWNHSHVSFVPLHYLNASNSGG 336  
Db 301 GYDMRKLNGTVSKHPLKAVTFVDNHDTPQGSLSTVQTWFKPLAYAFILTRREGSYPQ 360  
QY 337 YFDMRNILGVSQVQKHPITHAVTFVDNHDTPQSGALSFVQSWFKPLAYALILTRREGYPS 396  
Db 361 VFYDGMGTGKDSQREIPALKHKIEPILKARKQYAYCAQHDYFDHHDIVGWTREGDSSVA 420  
QY 397 VFYDGYGI--PTH-GVPSMKSIDPLQARQTYAGTQHDYFDHHDIIIGWTREGDSSHP 453



Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match						
1	376	21.2	5677	1	US-07-623-	Sequence 4,	Application 2,	2.72e-289
2	374	21.1	1777	1	US-08-146-	Sequence 33,	Application 1,	1.38e-287
3	374	21.1	1968	2	PCT-US94-0	Sequence 31,	Application 1,	1.38e-287
4	374	21.1	1968	2	PCT-US95-1	Sequence 31,	Application 1,	1.38e-287
5	374	21.1	2149	1	US-07-623-	Sequence 2,	Application 1,	1.38e-287
6	158	8.9	2182	3	5171673-5	Patent No. 5171673.		1.83e-105
7	98	5.5	7218	1	US-08-232-	Sequence 14,	Application 1,	1.31e-56
8	70	3.9	600	3	5171673-3	Patent No. 5171673.		1.24e-34
9	39	2.2	77	1	US-07-696-	Sequence 8,	Application 1,	7.90e-12
10	39	2.2	481	3	5171673-1	Patent No. 5171673.		7.90e-12
11	32	1.8	105	1	US-07-865-	Sequence 13,	Application 1,	3.80e-07

## ALIGNMENTS

## ALIGNMENTS

CC COUNTRY: USA  
CC ZIP: 94306  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible

***This Page Blank (uspto)***





***This Page Blank (uspto)***



***This Page Blank (uspto)***

QY 1333 gattttacggtgattactacggtat---acca-actcat-g-gtg---tcccttcgatg 1383  
 Db 1237 AAACACAAATTCACCGATCTTAAACGAGAAACAGTATGCGTACGAGACAGCAT 1296  
 QY 1384 aaatcaaatgatccactctgcagcgctcaacogtatgctcaggaaccaacat 1443  
 Db 1297 GATTATTCACCAACCATGACATGTCGGCTGACAAAGGAGGACGACAGCTCGGTGCA 1356  
 QY 1444 gattatttgatcatcatgatattatcggtgagcagagaaggagacagctccaccca 1503  
 Db 1357 AATTCAGGTTGGCGGCTTAATAACAGACGACCGCGTGGGGCAAGCAATGTATGTC 1416  
 QY 1504 aattcaggactgcaactattatgctcagatggccagggggttaataaatgatgtatgc 1563  
 Db 1417 GCGCGCAAAACCGCGGTGACATGGCATGACATTACCGGAACCGTTTCGGAGCCGGTT 1476  
 QY 1564 gggaaacataaagctggccaaagtatgagagatatcacgggaaataggctcgtgacogtc 1623  
 Db 1477 GTCATCAATTCCGAAGCGTCGGGAGAGTTTCACGTAACCGCGGGTCGTTTCAATTT 1534  
 QY 1624 accattaatcagatgggtgggggaatttcactgtaaacggagggcaggttcggtt 1681

RESULT 3  
 ID PCT-US94-01553A-31 STANDARD; DNA; UNC; 1968 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 31, Application PC/TUS9401553A  
 CC Sequence 31, Application PC/TUS9401553A  
 CC GENERAL INFORMATION:  
 CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
 CC TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
 CC NUMBER OF SEQUENCES: 68  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genencor International, Inc.  
 CC STREET: 180 Kimball Way  
 CC CITY: South San Francisco  
 CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US94/01553A  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Horn, Margaret A.  
 CC REGISTRATION NUMBER: 33,401  
 CC REFERENCE/DOCKET NUMBER: GC220-2  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 742-7536  
 CC TELEFAX: (415) 742-7217  
 CC INFORMATION FOR SEQ ID NO: 31:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 1968 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)

QY SEQUENCE 1968 BP; 573 A; 388 C; 509 G; 498 T; 0 OTHER.  
 Query Match 21.1%; Score 374; DB 2; Length 1968;  
 Best Local Similarity 66.1%; Pred. No. 1.38e-287;  
 Matches 951; Conservative 0; Mismatches 469; Indels 18; Gaps 9;  
 Db 260 AATGGGACCGCTGATGAGTATTGTAATGTACATGCCAATACGCGCCAACTTTGGAAG 319  
 QY 253 atggggacacatgatgcagctatttgaatggcattggccaaatgacggggaaccactggaac 312  
 Db 320 GCTTTCAAAACGACTCGGCATATTTCGGCTGAACACGCTATTACTGCGCTCTGGATTCCC 379  
 QY 313 aggttacgagatgacgcagctaaacttaagagtaagaggattaccgctgttggtattcct 372  
 Db 380 CCGGCTATATAGGAGACGAGCGAAGCGGATGTGGCTACGGCTGCTTACGACCTTTATGAT 439  
 QY 373 cctgcataagaggacttcgcaaaatgatgtgggtatggctcatgattgtacgat 432  
 Db 440 TTAGGGAGTTTCATCAAAAAGGAGCGTTTCGGCAAAAGTACGGCAAAAAGGAGAGCTG 499  
 QY 433 ctgtgtgagtttaacaaaagggaacogtccgtacaaaatatggcacaagagtcagttg 492  
 Db 500 CAATCTCCGATCAAAAAGTCTTCATTCCCGCGACATTAACGTTTACGGGGATGTGCTCATC 559  
 QY 493 caaggtgcogtgcacatcttgtaaaaataacgggattccaagtttatggggatgcgtgatg 552  
 Db 560 AACACACAAAGGCGGCTGATGCGACCGAAGATGTAACCGCGGTTGAAGTTCGATCCGCT 619  
 QY 553 aatcataaaggtggagcagcagcagagatggttaatgcggtggaagtgaaccgaagc 612  
 Db 620 GACCGCAACCGCGTAAATTCAGGAGACACCTTAATTAAGCCCTGCACACATTTTCTATTT 679  
 QY 613 aaccgaacccaagaatatcagggtgaatacacattgaagcatggaagaaatttgatttc 672  
 Db 680 CCGGGCGGGCGGACACATACAGCGATTTTAAATGGCATTTGTGTACCATTTTACGGAACC 739  
 QY 673 cctggaagaggaataaccattccaactttaaatggcgtggtatcatttggatggaca 732  
 Db 740 GATTGGACGAGTCCCGAAAGCT---GAACCGCATCTATAAGTTTC-A---A---GGAAG 790  
 QY 733 gattgggacatcagtcagctcagcttcagacaaaataataaattcagaggtaccggaag 792  
 Db 791 GCTTGGGATTGGGAAGTTTCCAATCAAAAACGGCAACTATGATTATTGTATGTCGCCAC 850  
 QY 793 gcatgggactgggaagttagatagagaacgcgaactatgattaccctatgtatgcagac 852  
 Db 851 ATCGATTATGACCATCTGATGTCGACGAGAAATTAAGAGATGGGCGACTTGTGTATGCC 910  
 QY 853 attgatattgatcatccagaagtaaatcaatgaacttagaaattgggagttgggtatata 912  
 Db 911 AATCAACTGCAATTGGACGGTTTCGCTTGTATGCTCTCAACACATTAATAATTTCTTTT 970  
 QY 913 aatacacttaactagatgatttagaatcgatgctgtgaacataataataacacagctat 972  
 Db 971 TTGCGGGATTGGGTTTATCATGTCAGGGAANAACGGGGAAGAAATGTTTACGTAGCT 1030  
 QY 973 acgagagatttgctaacacatgctgcgaaccacaggttaacccaatgtttgcagttgca 1032  
 Db 1031 GAATATGGCAGAAATGACTTGGCGCGCTGGAAACCTATTGGAACAAAACAAATTTAAT 1090  
 QY 1033 gaattttggaaaaaatgacctgtgcaatcgaaaaactatttaataaacaagttggaat 1092  
 Db 1091 CATTCAAGTCTTGGCGGCTTCATTATCAGTTTCCATCTGTCATCCACACAGGGAGCG 1150

***This Page Blank (uspto)***



***This Page Blank (uspto)***



Sep 16 10:24

US-08-952-741-1.mi

11

Db 911 AATGAACGCAATTGACGGTTCCTGTTGATGCTGTCGCAACACATTAATTTCTTTT 970  
 Qy 913 aatacaactaatctagatgattagaatcgatgctggaacataataataacagctat 972  
 Db 971 TTGCGGATTTGGTTTAATCATGTGAGGAAACCGGGAAGGAAATGTTTACGGTAGCT 1030  
 Qy 973 acgagagatggctaacacatgctgaacacacaggttaacacatgcttgcagattgca 1032  
 Db 1031 GAATATTCGCAATGACTTGGCGCGCTGCAAACTATTTGAACAAACAAATTTAAT 1090  
 Qy 1033 gaattttgaaaaatgacctgctgcaatcgaaacataatttaataaaaacagttggaat 1092  
 Db 1091 CATTGAGTGTTCGTCGTCGCTTCATTATCAGTTCATGCTGTCGACACAGGAGGC 1150  
 Qy 1093 cactcgtgttgatgcttctcttataattgtacaatgcatcctaagtgtggtgc 1152  
 Db 1151 GGCTATGATGAGGAAATTCCTGACGGTAGCGTCTTCCAGCATCGCTTGAATCG 1210  
 Qy 1153 tatattgatagaaatattttaaatggtctgtgacaaaacacocctatacatgca 1212  
 Db 1211 GTTACATTTGCTGAATACATATACAGCCGGGCAATCGCTTGAATGCTGTCGCA 1270  
 Qy 1213 gtcaactttgtgaacacatgactctcagcaggaagcattggaatccttggttcaa 1272  
 Db 1271 ACATGTTTAAAGCGTTCGTTACGCTTTTATCTCACAGGAACTCTGGATACCTCAG 1330  
 Qy 1273 tegtgttcaaacactggaatgcatgattgatctgacaaggagcaaggttaccttcc 1332  
 Db 1331 GTTTCTACGGGATATCTACGGGACGAAGAGACTCCAGCGCGAAATTCCTGCCCTG 1390  
 Qy 1333 gtatttacggtgattactaaggtat----acca-actaat-g-gtg----ttccttgatg 1393  
 Db 1391 AAACACAAAATTAACCGATCTTAAAGCGAGAAACAGTATCGTACGAGACACAGCAT 1450  
 Qy 1384 aaatctaaatgatccactcttgcaggcagctcaaacgtatgctcaggaacccaacat 1443  
 Db 1451 GATTATTCGACACCATGATGCTGGCTGGACAGGAGGACAGCTCGGTTGCA 1510  
 Qy 1444 gattattttgatcatcatgatattatcgctggacgagagaaggggacagctccacca 1503  
 Db 1511 AATTGAGTTTGGCGCAATTAACACAGCAGCCGCTGGCGCAAGCAATGATGTC 1570  
 Qy 1504 aattcagacttgcaactattatgctcgatggccaggggttaataaatggatgatgct 1563  
 Db 1571 GCGCGGCAAAACCGCGGTGACATGGCATGCAATTCGCGAAACCGTTGCGACCGGTT 1630  
 Qy 1564 gggaaacataaagctggcaagatgagagatatacccggaataaggtctggtaccgtc 1623  
 Db 1631 GTCATCAATTCGGAAGCGCTGGGAGGTTTTCAGCTAAACGGCGGCTCGGTTCAATT 1688  
 Qy 1624 accattaatgcagatggttgggggaatttcactgtaaacgaggggcggtttcggtt 1681

RESULT 5

ID US-07-623-953-2 STANDARD; DNA; UNC; 2149 BP.

AC xxxxxx

DE Sequence 2, Application US/07623953

CC Sequence 2, Application US/07623953

CC Patent No. 5364782

CC GENERAL INFORMATION:

CC APPLICANT: Quax, Wilhelmus J

CC APPLICANT: Laroche, Yves

Sep 16 10:24

US-08-952-741-1.mi

12

CC APPLICANT: Vollebregt, Adrianus W.H.  
 CC APPLICANT: Stanseene, Patrick  
 CC APPLICANT: Lauwereys, Marc  
 CC TITLE OF INVENTION: MUTANT MICROBIAL ALPHA-AMYLASES WITH  
 CC TITLE OF INVENTION: INCREASED THERMAL, ACID AND/OR ALKALINE STABILITY  
 CC NUMBER OF SEQUENCES: 7  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM  
 CC STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR  
 CC CITY: PALO ALTO  
 CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94306  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/623,953  
 CC FILING DATE: 19901129  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Rae-Venter, Barbara  
 CC REGISTRATION NUMBER: 32,750  
 CC REFERENCE/DOCKET NUMBER: GBRO-025/0005  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415-494-7622  
 CC TELEFAX: 415-857-0663  
 CC TELEX: 380816 COOLEY PA  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 2149 base pairs  
 CC TYPE: NUCLEIC ACID  
 CC STRANDEDNESS: double  
 CC TOPOLOGY: linear  
 CC FEATURE:  
 CC NAME/KEY: CDS  
 CC LOCATION: 382..1920  
 CC FEATURE:  
 CC NAME/KEY: mat peptide  
 CC LOCATION: 469..1920  
 CC SEQUENCE 2149 BP; 621 A; 420 C; 559 G; 549 T; 0 OTHER.

Query Match 21.1%; Score 374; DB 1; Length 2149;  
 Best Local Similarity 66.1%; Pred. No. 1.38e-287;  
 Matches 951; Conservative 0; Mismatches 469; Indels 18; Gaps 9;

Db 478 AATGGGACGCTGATGACGATATTTTGAATGCTACATGCCAATGACGGCAACATTCGAG 537

Qy 253 aatgggaccatgatgacgtattttgaatgacatttgccaaatgacgggaaccactggaac 312

Db 538 CGTTTCCAAAACGACTCGGCATATTTTGGCTGAACACGCGTATTACTGCGCTTGGATTCCC 597

Qy 313 aggtacagatgacgcagctaaacttaagagataaaggattaccgctgtttgattcct 372

Db 598 CCGGCATATAAGGCAACGACGACGATGTGGCTACGGTACGGCTTTATGAT 657

Qy 373 cctgcatggaaggaggacttcgcgaataatgatgttggtatggtgcctatgatttgcagat 432

Db 658 TTAGGGGAGTTTCATCAAAAAGGGGCGTTTCGGACAAAGTACGGCACAAAAGGAGAGCTG 717

Qy 433 cttggtgagtttaacccaaaagggaacgcgtccgtacaaaatattggcacaaggagtcagttg 492

***This Page Blank (uspto)***

Db 718 CAATCTGCATCAAAAGTCTTCATTCCTCCGCGACATTAACGTTTACGGGATGTGTCATC 777  
 Qy 493 caagggtgcgtgacatctcttgaataataaacggaattcaagtttatgggatgctcgtgatg 552  
 Db 778 AACCAAAAGCGCGCTGATCGCCGAGAGATGATTAACCGCGGTTGAGTGCATCCCGCT 837  
 Qy 553 aatcataaagtgagacagcgagacagagatggttaaatgcggtggaagtgaaccgaagc 612  
 Db 838 GACCGAACCCGCTTAATTTGAGGAGAACACCTTAATTAAGCCCTGCACATTTTCATTTT 897  
 Qy 613 aaccgaaccaagaataatcaagtggaatacaccatggaagcatgagcgaatttgatttc 672  
 Db 898 CCGGCGCGCGAGACACATACAGCGATTTTAAATGGCATTTGCGATTTTACGGGAACC 957  
 Qy 673 cctggaagagaataaccattccacttaaatgagcgtggtatcaatttgatggagaca 732  
 Db 958 GATTGGAGCAGTCCGAAAGCT---GAACCGCATCTATAAGTTTC-A--A--GGAAG 1008  
 Qy 733 gattggatcagtcacgtcagcttcagagaacaaatataataattcagaggtaccgggaag 792  
 Db 1009 GCTTGGGATTGGGAGTTCCTCAATGAAAGCGCAACTATGATTAATTTGATGTATGCGGAC 1068  
 Qy 793 gcatggactgggaagatagatagagaacggaactatgattaccttatgtatgcagac 852  
 Db 1069 ATGATTAATGACATCCGTGATGCGCAGCAGAAATTAAGAGATGGCGACTGGTATGCC 1128  
 Qy 853 attgataggatcatccagagaataatcaatgaacttagaattggagttggtatataca 912  
 Db 1129 AATGAACATGCAATGAGCGGTTCCGCTTCTGATGCTGTCACACATTAATTTCTTTT 1188  
 Qy 913 aatacactaatctagaggtattagaatcgatgctggaacataattaataacagctat 972  
 Db 1189 TTCCGGGATTGGGTAAATGTCAGGGAAGAAACCGGGAAGAAATGTTTACGGTAGCT 1248  
 Qy 973 acgagagattggtcaacacatgctgaacacacaggttaacacacatggttcgagttgca 1032  
 Db 1249 GAATATTGGCAAAATGACATGCGCGGCTGGAAACTATTTGCAAAAACAAATTTTAAAT 1308  
 Qy 1033 gaattttgaaataatgacctgctgaatcgaaactatttaataaacaagaattgggaat 1092  
 Db 1309 CATTGAGTGTTCAGTCGCGCTTCATTATGATGTCATGCTGCTGCAACAGTCCGTTGAAATCG 1368  
 Qy 1093 cactcgtgttcgatgttcctctcttataataattgtacaatgcatctaatagtggtgac 1152  
 Db 1369 GGCTATGATATGAGGAATTCCTGACGCTAGCGTCTTCCACAGATCCGTTGAAATCG 1428  
 Qy 1153 tattttgatagagaataattttaaattggtctgctgacaaaacacctatatacatgca 1212  
 Db 1429 GTTACATTTGTCGATTAACCATATACAGCGCGGGCAATCGCTTGATGTCGATGTCGCA 1488  
 Qy 1213 gtacattttgtgaataaccatgactctcagccaggaagacattggaactcttggttcaa 1272  
 Db 1489 ACATGTTTAAAGCGGTTGCTTACGCTTTTATTCACAAGGGAATTCGATACCCCTCAG 1548  
 Qy 1273 tcdtgggtcaaacaccactggcatatgattgattctgacaagggaggaaggttacccttcc 1332  
 Db 1549 GTTTTCTACGGGATATGATGCGGAGCAAGGAGCTCCCGCGCGAATTCCTGCCTTG 1608  
 Qy 1333 gtattttcaggtgattactacggtat----acca-actcat-g-gtg----tcccttgatg 1383  
 Db 1609 AAACACAAATTCGACCGATCTTAAGAGCGAAGAAACAGTATGCTGACGACACACAT 1668  
 Qy 1384 aaatctaaattgatccactctgcagggcagtcgaacgtatgctcaacggaacccaacat 1443

Db 1669 GATTATTTCACCATGATCTCGCTGCAGACAGGGAAGGCGACAGCTCGGTTGCA 1728  
 Qy 1444 gattattttgatcatcatgataattcgctgagcagagagaadagggagcagctcccacca 1503  
 Db 1729 AATTGAGTTTCGGCGCATTAATTAACAGCGGACCCCGTGGGCAAGCGAATGTATGTC 1788  
 Qy 1504 aattcaggacttgcaactattatgtccgatggccagggggttaataaatggatgtatgctc 1563  
 Db 1789 GCGCGGCAAAAGCGGTGAGACATGCGATGACATTAACCGGAACCGTTCGGAGCGGTT 1848  
 Qy 1564 ggaataacataaagctggccaagtatgagagatatcaccggaataatgctgtaccgctc 1623  
 Db 1849 GTTCATCAATTCGGAAGCTGGGAGAGTTTTCACGTAAACCGGGTTCGTTTCAATTT 1906  
 Qy 1624 accattaatgcagatggttcgggggaatttcaactgtaaacgaggggcagtttcggttt 1681  
 RESULT 6  
 ID 5171673-5 STANDARD; DNA; UNC; 2364 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673.  
 CC Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:5:  
 CC LENGTH: 2182  
 SQ Sequence 2364 BP; 611 A; 482 C; 528 G; 561 T; 182 other;  
 Query Match 8.9%; Score 158; DB 3; Length 2182;  
 Best Local Similarity 58.7%; Pred. No. 1.83e-105;  
 Matches 651; Conservative 0; Mismatches 451; Indels 7; Gaps 7;  
 Db 376 GAATCATFACATCATGCGCTTTTGAATGGAATACCCAGCAGACGCGCATTTGGA 435  
 Qy 252 gaatggaccatgatgcagtattttgaatggcatttgccaataatgacgggaaccactgaa 311  
 Db 436 CCGGCTCAAGAAATGGCGCTGAATTAAGAAAGCGGATTTGATGCGCTCTGGGTTCC 495  
 Qy 312 caggttacgagatgacgcagctaaccttaagagataaaggattaccgctgtttggattcc 371  
 Db 496 CCCGCTCAAAAAGCAGCTCAGCATGGAATGTTACGGGCTGTACGACCATTTACGA 555  
 Qy 372 tctcgatggaaggggacttcgcaaaatgatgtggtgatggtgcctatgatttgaaga 431  
 Db 556 CCGCGGAGTTTACCAAGAAAGCGCGCTCAGGACAAAGTACGGGACAAACAGCAATT 615  
 Qy 432 tcttggtgagtttaaccaaagggaaccgtccgtacaaaataatggcacaagagtcagt 491  
 Db 616 ACATGAAGCATCAACCGCTGCGCAGCAGCATATCCAGGTCTATATCGATCTGCTAT 675  
 Qy 492 gcaaggtgcgctgacatctttgaaaaataaacgggatccaagtttatgggagatgcgtgat 551  
 Db 676 GAACCATTAAGCGCGCGCGATGAACCCGAATCTTTCCAAAGTGTGGAGGTTCGACCCGAT 735  
 Qy 552 gaatcataaaggtgagcagcgagcgagagatggttaaatgcggtggaagtgaaccgaag 611  
 Db 736 GGACCGCAACAAAGAAATTTCCGAAACCGTTTGAATAGAGGCTGGACAAAGTTCAATTT 795

***This Page Blank (uspto)***

[illegible]

RESULT	7
ID	US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC	xxxxxx
DT	
DE	Sequence 14, Application US/08232463
CC	Sequence 14, Application US/08232463
CC	Patent No. 5670367
CC	GENERAL INFORMATION:
CC	APPLICANT: DORNER, F.
CC	APPLICANT: SCHEIFLINGER, F.
CC	APPLICANT: FALKNER, F. G.
CC	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

```
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22131-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZapt-Fla
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match          5.5%; Score 98; DB 1; Length 7218;
Best Local Similarity 1.1%; Pred. No. 1.31e-56;
Matches      4; Conservative    227; Mismatches 133; Indels   0; Gaps   0;

Db 1075 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1134
      :: : ::::: :: : ::::: :: : : ::::: :: : : ::::: :: : :
Cp 827 ttgcgttctatatactaccagtcgccatgcttcgcgtacctgaatttat 768

Db 1135 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1194
      ::::: ::::: :: : : ::::: :: : : ::::: :: : : :::::
Cp 767 atttgttgaaagcagcgactgatcccaactgtcccacaaaatgataccagcg 708

Db 1195 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1254
      : :: : : : : : : : : : : : : : : : : : : : : : : : :
Cp 707 catttaagtggaaagggtattctcttcgccagggaataaatctgcctcatgctca 648

Db 1255 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1314
      : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 647 atgggtattcacctgatattcttggttcggttgcttcggttcacttccaacgcat 588

Db 1315 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1374
      : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 587 accatctcttcgcgtctgcacactttatgattcatcacgacatccccataaactga 528
```

***This Page Blank (uspto)***

Db 1375 YY 1434  
 Cp 527 atccggtatttttcaagagatgcacggcaccttgcaactgactccttgatgcatattt 468  
 Db 1435 GTAC 1438  
 Cp 467 gtac 464

RESULT 8

ID 5171673-3 STANDARD; DNA; UNC; 650 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:3:  
 CC LENGTH: 600  
 CC Sequence 650 BP; 189 A; 124 C; 135 G; 152 T; 50 other;

Query Match 3.9%; Score 70; DB 3; Length 600;  
 Best Local Similarity 65.2%; Pred. No. 1.24e-34;  
 Matches 150; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Db 315 GAATCATACATCATGCGATTTTTCATGGAATACGCCAGCAGCGGCATTCGAA 374  
 Qy 252 gaatggaccatgatgcagtattttgaatggcatttgcacaaatgacgggaaccactgaa 311  
 Db 375 CCGGCTGAAGAAATGGCGCCCTGAATTAAGAAAGCGGATTCGCGCTGCTCC 434  
 Qy 312 caggttaagatgacgacgtactaaagatgaaaggtattcgcgtgttgattcc 371  
 Db 435 CCGGCTGACAAAGACATGACATGACATGTTACGGGGTGTACCAATTCACGA 494  
 Qy 372 tctgcatggaagggaacttcgcaaatgatgttggtatggtcctatgattgtacga 431  
 Db 495 CTCTGGGCGATTTCACCAAGAGCCCTCAGCAGCAAAAGTACGGGACAA 544  
 Qy 432 tcttggtgagtttaaccaaaagggaacgcgtccgtacaaaatattggcaaa 481

RESULT 9

ID US-07-696-551B-8 STANDARD; DNA; UNC; 77 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 8, Application US/07/696551B  
 DE Sequence 8, Application US/07/696551B  
 CC Patent No. 5232841  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Hashimoto, Tamotsu  
 CC APPLICANT: Tsujimura, Atsushi  
 CC APPLICANT: Ueda, Shigezo  
 CC TITLE OF INVENTION: Process for Preparing Peptide  
 CC NUMBER OF SEQUENCES: 12  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 CC ADDRESSEE: Dunner

CC STREET: 1300 I Street, N.W., Suite 700  
 CC CITY: Washington  
 CC STATE: D.C.  
 CC COUNTRY: USA  
 CC ZIP: 20005-3315  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: MS-DOS/PC-DOS  
 CC SOFTWARE: Patent in Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/696,551B  
 CC FILING DATE: 19910509  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 2-122166  
 CC FILING DATE: 11-MAY-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 2-334575  
 CC FILING DATE: 30-NOV-1990  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Lawrence M. Lavin, Jr.  
 CC REGISTRATION NUMBER: 30,768  
 CC REFERENCE/DOCKET NUMBER: 2481-1070  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (202) 408-4000  
 CC TELEFAX: (202) 408-4400  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 77 bases pairs  
 CC TYPE: NUCLEIC ACID  
 CC STRANDEDNESS: both  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)  
 CC SEQUENCE 77 BP; 22 A; 16 C; 21 G; 18 T; 0 OTHER.

Query Match 2.2%; Score 39; DB 1; Length 77;  
 Best Local Similarity 83.1%; Pred. No. 7.90e-12;  
 Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 17 AATGGACCTGATGCTGATTTTGAATGCTACATGCCAATGACGGCCAACTTGGAA 75  
 Qy 253 aatggaccatgatgcagtattttgaatggtcatttgcacaaatgacgggaaccactggaa 311

RESULT 10

ID 5171673-1 STANDARD; DNA; UNC; 521 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673  
 CC Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:1:  
 CC LENGTH: 481  
 CC Sequence 521 BP; 145 A; 102 C; 96 G; 138 T; 40 other;  
 Query Match 2.2%; Score 39; DB 3; Length 481;

***This Page Blank (uspto)***



```

Best Local Similarity 68.9%; Pred. No. 7,90e-12;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Db 376 GAATCATCAATCATGACGATTTTTTGAATGGAATACCGCAGACAGCGGACGACATTGGAA 435
Qy 252 gaatgggaccatgatgcagtatatttgatggcatttgcacaatgacgggaaccacggaa 311

Db 436 CGCGCTCAAGAAATGCGCCCTGATTAAGAAAGAAAGCGCGATT 478
Qy 312 caggttacgagatgacgcagctaaccttaagagtagtaaggatt 354

RESULT 11
ID US-07-865-662F-13 STANDARD; DNA; UNC; 105 BP.
AC xxxxxx
DT
DE Sequence 13, Application US/07865662F
CC Sequence 13, Application US/07865662F
CC Patent No. 5451670
CC GENERAL INFORMATION:
CC APPLICANT: Marcia M. Miller
CC TITLE OF INVENTION: Restriction Fragment Length
CC TITLE OF INVENTION: Polymorphism Test For Haplotyping Domesticated Fowl
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: City of Hope
CC STREET: 1500 East Duarte Road
CC CITY: Duarte
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 91010-0269
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3M Double Density 5 1/4" diskette
CC COMPUTER: Wang PC
CC OPERATING SYSTEM: MS DOS Version 3.20
CC SOFTWARE: Microsoft
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/865,662F
CC FILING DATE: 07 April, 1992
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/688,326
CC FILING DATE: 22 April 1991
CC APPLICATION NUMBER: 07/588,922
CC FILING DATE: 27 September 1990
CC APPLICATION NUMBER: 07/210,405
CC FILING DATE: 23 June 1988
CC APPLICATION NUMBER: US 07/130,529
CC FILING DATE: 9 December 1987
CC APPLICATION NUMBER: US 07/068,176
CC FILING DATE: 30 June 1987
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Irons, Edward S.
CC REGISTRATION NUMBER: 16,541
CC REFERENCE/DOCKET NUMBER: No. 5451670e
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 783-6040
CC TELEFAX: (202) 783-6031
CC TELEX: No. 5451670e
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 105
CC TYPE: Nucleic Acid
CC STRANDEDNESS: Double

```

```
CC      TOPOLOGY: Linear
CC      MOLECULE TYPE: DNA
CC      ORIGINAL SOURCE: Synthetically Prepared
CC      IMMEDIATE SOURCE: Synthetically Prepared
SQ      SEQUENCE 105 BP; 15 A; 0 G; 8 C; 1 T; 81 OTHER.

Query Match          1.8%; Score 32; DB 1; Length 105;
Best Local Similarity 16.5%; Pred.No. 3.80e-07;
Matches 15; Conservative 28; Mismatches 48; Indels 0; Gaps 0;

Db       9 AKSNNNKSNNAVKGNNAKSNGKNKSANAVNNGNNAKSNNNKSNMNGTRDNKGNNRNA 68
Qy       976 agagattgctaacacatgtcgtaaccacccagcgtaaaccaaatgttgcagtgcagaa 1035
         |: : :: |: : |: : |: : | : | : | : | : | : | : | : | : | : | : |
Db       69 KNNNSGVADNKNNANNYDNGSGVADNKNA 99
Qy       1036 ttttgaaaaatgacctgtgctgaactgcaga 1066
         : :|:|: |: : : : |: : |: |

RESULT 12
ID      US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
DT      ACACxxxxxx
DE      Sequence 14, Application US/08232463
CC      Sequence 14, Application US/08232463
CC      Patent No. 5670367
CC      GENERAL INFORMATION:
CC      APPLICANT: DORNER, F.
CC      APPLICANT: SCHEIFLINGER, F.
CC      APPLICANT: FALKNER, F. G.
CC      TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
CC      NUMBER OF SEQUENCES: 52
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Foley & Lardner
CC      STREET: 1800 Diagonal Road, Suite 500
CC      CITY: Alexandria
CC      STATE: VA
CC      COUNTRY: USA
CC      ZIP: 22313-0299
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/232,463
CC      FILING DATE:
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US/07/935,313
CC      FILING DATE:
CC      APPLICATION NUMBER: EP 91 114 300.6
CC      FILING DATE: 26-AUG-1991
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: BENT, Stephen A.
CC      REGISTRATION NUMBER: 29,768
CC      REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (703)836-9300
CC      TELEFAX: (703)683-4109
CC      TELEX: 899149
CC      INFORMATION FOR SEQ ID NO: 14:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 7218 base pairs
```

***This Page Blank (uspto)***

CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC IMMEDIATE SOURCE:  
CC CLONE: pTZgpt-fls  
SQ SEQUENCE 7218 BF; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

[illegible]

RESULT	13	
ID	US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.	
AC	xxxxxx	
DT		
DE	Sequence 5, Application US/08238163	
CC	Sequence 5, Application US/08238163	
CC	Patent No. 5569830	
CC	GENERAL INFORMATION:	
CC	APPLICANT: BENNETT, Alan	
CC	APPLICANT: LABAVITCH, John M.	
CC	APPLICANT: POWELL, Ann	
CC	APPLICANT: STOTZ, Henrik	
CC	TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL	
CC	TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE	
CC	NUMBER OF SEQUENCES: 24	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Townsend and Townsend Khourie and Crew	
CC	STREET: Stewart Street Tower, One Market Plaza	
CC	CITY: San Francisco	
CC	STATE: California	
CC	COUNTRY: US	
CC	ZIP: 94105-1493	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	

```

CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc feature
CC LOCATION: 1..215
CC
CC OTHER INFORMATION: /standard name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
CC
CC SEQUENCE 215 BP: 15 A: 8 G: 25 C: 26 T: 141 OTHER:
CC

```

Query Match	1.6%	Score 29;	DB 1;	Length 215;
Best Local Similarity	19.4%;	Pred. No. 3.09e-05;		
Matches	28;	Conservative 52;	Mismatches 63;	Indels 1; Gaps 1;
Db	8	SSSVSRVASCNDKAKKDGNTTSSWTTDCNRITGVCVDITTYRVNNDSGHNKYSANYN	67	
	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::
Qy	458	ccgtcgctacaaaataatggcacaaaggagtcagttgcgaaggtgcgtgacatctttgaaaa	517	
Db	68	YGCNNVGAAKTHYYTHTNVSGADSKVTVDPSYNASGTSSSNGCTGDCNRASGADSYG-SSKTA	126	
	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::
Qy	518	ataacgggattcaagtttatgggtggtcgctgatcataaagaaggagcagacggga	577	
Db	127	MTSPNETGKTANNVDSRNMGDAS	150	
	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::
Qy	578	cadaagatgataaatccggtgaaq	601	

RESULT 14  
ID US-08-238-163-5 STANDARD; DNA: UNC; 215 BP.

DI	Sequence 5, Application US/08238163
DE	Sequence 5, Application US/08238163
CC	Patent No. 5569830
CC	GENERAL INFORMATION:
CC	APPLICANT: BENNETT, Alan
CC	APPLICANT: LABAVITCH, John M.
CC	APPLICANT: POWELL, Ann
CC	APPLICANT: STOTZ, Henrik
CC	TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC	TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC	NUMBER OF SEQUENCES: 24
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Townsend and Khourie and Crew
CC	STREET: Steuart Street Tower, One Market Plaza
CC	CITY: San Francisco
CC	

**This Page Blank (use)**

Sep 16 10:24

US-08-952-741-1 mi

23

CC STATE: California  
 CC COUNTRY: US  
 CC ZIP: 94105-1493  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/238,163  
 CC FILING DATE: 03-MAY-1994  
 CC CLASSIFICATION: 800  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Bastian, Kevin L.  
 CC REGISTRATION NUMBER: 34,774  
 CC REFERENCE/DOCKET NUMBER: 2307E-540  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 543-9600  
 CC TELEFAX: (415) 543-5043  
 CC INFORMATION FOR SEQ ID NO: 5:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 215 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: unknown  
 CC MOLECULE TYPE: protein  
 CC FEATURE:  
 CC NAME/KEY: misc feature  
 CC LOCATION: 1..215  
 CC OTHER INFORMATION: /standard name= "Deduced amino acid  
 CC OTHER INFORMATION: sequence of PCIP from bean."

SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.  
 Query Match 1.5%; Score 26; DB 1; Length 215;  
 Best Local Similarity 13.7%; Pred. No. 2.12e-03;  
 Matches 23; Conservative 61; Mismatches 84; Indels 0; Gaps 0;

Db 34 TDCQRTMGVCDPTTYRVNNDGSHKYSANVYGNVGAATKTHYTHINVSADSKT 93  
 Cp 1091 ttccaaactgttttaataatgttttcgattgcagaggtcattttccaaaattct 1032  
 Db 94 VTDSTNAGTSSNGGDRSGADSYGSKTAMTSRRTGKTANNVDSRNMCDASVGS 153  
 Cp 1031 gcaactgcaaacatggtttacctgtggttacgcacatgtgttagccaatctctcgt 972  
 Db 154 DRNITKHKNSADGKVGSKNGGNRRNRYGTGTSNVSNCGGNKRDV 201  
 Cp 971 tagctgtatttaataatgtttcacagcatcgatcttaaatccatctaga 924

RESULT 15  
 ID PCT-US95-10426-3 STANDARD; DNA; UNC; 34 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 3, Application PC/TUS9510426  
 CC Sequence 3, Application PC/TUS9510426  
 CC GENERAL INFORMATION:  
 CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
 CC TITLE OF INVENTION: An Improved Cleaning Composition  
 CC NUMBER OF SEQUENCES: 68  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genencor International  
 CC STREET: 180 Kimball Way  
 CC CITY: South San Francisco

Sep 16 10:24

US-08-952-741-1 mi

24

CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US95/10426  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: STONE, Christopher L.  
 CC REGISTRATION NUMBER: 33,401  
 CC REFERENCE/DOCKET NUMBER: GC220-3  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 742-7536  
 CC TELEFAX: (415) 742-7217  
 CC INFORMATION FOR SEQ ID NO: 3:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 34 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)  
 CC SEQUENCE 34 BP; 9 A; 7 C; 8 G; 10 T; 0 OTHER.

Query Match 1.4%; Score 24; DB 2; Length 34;  
 Best Local Similarity 85.3%; Pred. No. 3.17e-02;  
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 TCATGCAGTACTTGTGAATGGTACCTGCCCAATGA 34  
 Qy 263 tgatgcagattttgaatggcatttgccaaatga 296

Search completed: Wed Sep 16 10:27:47 1998  
 Job time : 98 secs.

**This Page Blank (usp)**

\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Sep 16 08:46:42 1998; MasPar time 2402.17 Seconds  
Tabular output not generated. 1358.207 Million cell updates/sec

Title: >US-08-952-741-1  
Description: (1-1776) from US08952741.seq

Perfect Score: 1776  
N.A. Sequence: 1 atataaattgaaatgaaca.....ccaataaattggaagctt 1776  
Comp: tatattaaactttactgt.....ggttatattaaaccttcgaa

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155

Database: 1:em\_ba 2:em\_htg 3:em\_huml 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro  
genbank107  
12:gb\_ba 13:gb\_htg 14:gb\_in 15:gb\_om 16:gb\_ov 17:gb\_pat  
18:gb\_ph 19:gb\_pr1 20:gb\_pr2 21:gb\_ro 22:gb\_ro 23:gb\_st  
24:gb\_sts 25:gb\_sy 26:gb\_un 27:gb\_vi

Statistics: Mean 11.347; Variance 5.064; scale 2.241

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	785	44.2	2397	12	BACAMYG6	Bacillus sp. (alkaloph
2	473	26.6	2447	12	BSU22045	Bacillus sp. alpha-amy
3	432	24.3	1990	12	AF032864	Bacillus stearothermop
4	432	24.3	3048	12	BACAMYLB	Bacillus stearothermop
5	430	24.2	1680	17	E01181	DNA encoding highly th
6	430	24.2	1719	17	E01180	DNA encoding highly th
7	430	24.2	2393	12	BSU75445	Bacillus sp. MK 715 a
8	428	24.1	1814	12	BSAMYSEN	B.stearothermophilus D
9	428	24.1	1891	12	BACAMYABS	B.stearothermophilus a
10	428	24.1	2169	12	BSAMYLAI	Bacillus stearothermop
11	424	23.9	1650	17	E01157	DNA sequence of B.stea
12	424	23.9	2066	12	BACAMYS	B.stearothermophilus a
13	376	21.2	1539	17	A47677	Sequence 1 from Patent
14	376	21.2	5877	17	A21895	Nucleotide sequence of
15	374	21.1	1449	17	A23402	B.licheniformis gene f

16	374	21.1	1449	17	A27772	Alpha amylase coding s	0.00e+00
17	374	21.1	1777	17	A17930	Alpha amylase gene.	0.00e+00
18	374	21.1	1777	17	A24553	Sequence 33 from paten	0.00e+00
19	374	21.1	1948	12	BACAMYLG	Bacillus licheniformis	0.00e+00
20	374	21.1	1948	12	BACAMYABL	B.licheniformis alpha-	0.00e+00
21	374	21.1	1968	12	BACAMYS	B.licheniformis amylase	0.00e+00
22	374	21.1	2149	17	A21893	Nucleotide sequence of	0.00e+00
23	374	21.1	2829	17	I08486	Sequence 13 from Paten	0.00e+00
24	374	21.1	4190	17	I08488	Sequence 19 from Paten	0.00e+00
25	355	20.0	1536	17	E01158	DNA sequence of B.lich	1.03e-293
26	352	19.8	1972	17	A20154	alpha-amylase gene (an	6.99e-291
27	352	19.8	2084	12	BACAM	Bacillus amyloliquefac	6.99e-291
28	158	8.9	2182	17	I08183	Sequence 1 from Patent	1.48e-110
29	156	8.8	1900	12	AF055987	Streptococcus mutans i	9.56e-109
30	141	7.9	2050	12	BCAMYE	B.circulans amylase	3.23e-95
31	128	7.2	2302	12	SB004956	Streptococcus bovis am	1.38e-83
32	114	6.4	2300	12	AB000830	Streptococcus bovis ge	3.47e-71
33	98	5.5	7218	17	I66494	Sequence 14 from paten	3.27e-57
34	84	4.7	454	12	BACAMYL	B.licheniformis alpha-	3.27e-45
35	84	4.7	454	17	I00838	Sequence 10 from Paten	3.27e-45
36	83	4.7	576	17	A00157	Nucleotide sequence 3	2.30e-44
37	83	4.7	576	17	A00604	B.amyloliquefaciens al	2.30e-44
38	83	4.7	576	17	A00155	Nucleotide sequence 1	2.30e-44
39	83	4.7	576	17	A00158	Nucleotide sequence 4	2.30e-44
40	83	4.7	576	17	A00605	B.amyloliquefaciens al	2.30e-44
41	83	4.7	576	17	A00156	Nucleotide sequence 2	2.30e-44
42	65	3.7	465	17	E00040	DNA coding of alpha-am	2.01e-29
43	65	3.7	537	12	BAAMYL	Bacillus amyloliquefac	2.01e-29
44	63	3.5	465	17	A00607	B.amyloliquefaciens al	8.34e-28
45	63	3.5	465	17	A00608	B.amyloliquefaciens al	8.34e-28

ALIGNMENTS

RESULT 1  
LOCUS BACAMYG6 2397 bp DNA BCT 15-MAR-1989  
DEFINITION Bacillus sp. (alkalophilic) G6-amylase gene, complete cds.  
ACCESSION M18862  
NID g142496  
KEYWORDS amylase; maltohexaoase-producing amylase.  
SOURCE Bacillus sp. (alkalophilic strain #707) DNA, clone pTUE306.  
ORGANISM Eubacteria; Firmicutes; Low G+C gram-positive bacteria;  
Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 2397)  
AUTHORS Tsukamoto,A., Kimura,K., Ishii,Y., Takano,T. and Yamane,K.  
TITLE Nucleotide sequence of the maltohexaoase-producing amylase gene from  
an alkalophilic Bacillus sp. #707 and structural similarity to  
liquefying type alpha-amylases  
JOURNAL Biochem. Biophys. Res. Commun. 151, 25-31 (1988)  
MEDLINE 88162814  
COMMENT Draft entry and computer-readable sequence for [1] kindly provided  
by K.Yamane, 02-JUN-1988.  
FEATURES  
source  
1. .2397  
/organism="Bacillus sp."  
/db\_xref="taxon:1409"  
729..827  
/note="G6-amylase signal peptide"  
729..2285  
/note="G6-amylase precursor"  
/codon\_start=1  
/transl\_table=11  
/db\_xref="PID:g142497"  
/translation="MKMRTQKGLSLLAFLLVITIPFTLVDEAHNNTGTFMMQ  
YFWMYLPNDGNHNWNRNSDASLKSIGTAVIPPAWKAGSQNDVGAYDLYDIGEF  
NQGTGRTYRYSQQAAYTSLKNGIQYGVVMHKGADATMVAEYVNPNNR  
NQVTEGTYTEATWTRDFPCRGTHSFQKRWYHFDGVDQDSRLNNRIYKFGHGRK  
ANDWEVDTEGNYDLYMADIDMDHPVNLNMGWYNTLGLDGFRIADVAHKY  
SFRDNNHRSATGNMFAVAEKNLDGAIENLYLQKTNHNSVDFVPLHYNLAS  
KSGNVDMMRNIFGTGVVQVHSHAVTFVDNHDSDQEALESFVEWFKPLAYLTLR  
EQGYPVFYGDYGTIPTHGVPMARKIDPILKQYKAYQNDYLDHNTIGTREG  
NTAHPNSGLATIMSDGAGSKWFMVGRKAGQVMSDITGRTGTVTINADGWNFSVN

```
GGSVSIWVVK"
mat_peptide      828..2282
                  /note="G6-amylose"
BASE COUNT      788 a      346 c      549 g      714 t
ORIGIN          1 bp upstream of BamHI site.

Query Match      44.28; Score 785; DB 12; Length 2397;
Best Local Similarity 76.7%; Pred. No. 0.00e+00;
Matches 1128; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

Db 823 AACGACATCATACGTCACCAAGCGGACGATGATGCAATCTTTGAATGTATCTACTA 882
Qy 233 aagccataataatggagcaatggaccatgatcgatgatttttgaatggcattgcaaa 292

Db 883 ATGACGAAATCATTTGAAATCGATTAACCTCTGTATGCGAGTAACCTTTAAAGCAAGGA 942
Qy 293 atgacgggaacctggaaacaggttaacagatgacgcagcttaacattaaagtaaaagga 352

Db 943 TTACAGGGGTGGATTCCTCCAGCATGGAAGGCGCTTCTCAAAATGACGTAGATACG 1002
Qy 353 ttaccgctgttgattccctcctgcatgggaaggggacttcgcaaaatgatgtgggtatg 412

Db 1003 GAGCCTATGACCTGTATGATCTGGGAGAAATTAATCAAAAAGGTACCGTCGTACAAAT 1062
Qy 413 gtgcctatgatgtgtcagatcctgttgagtttaacccaaaggaacccgtcgtacaaaat 472

Db 1063 ATGGAACACGTCAGTTTCAAGCTGCGGTAACCTCTTTAAATAATATGGAATTCAG 1122
Qy 473 atggcaagagtgatgtgcaaggtgcctgacatcttggaaaaataacgggattcaag 532

Db 1123 TATATGGTGACGTGTTTATGAATACAAAGGTGGCGCAGACGCTACTGAAATGTAAGGG 1182
Qy 533 ttatggggatgctgtgatgaatacaaaaggtggagcagcgggacagagatggtaaatg 592

Db 1183 CGCTTCAAGTGAATCCCAATACCGTAACCAAGNAGTACTGTGTATATACCAATTCAG 1242
Qy 593 cgggtgaagtgaaccgaagcaacgaacaaataatcaggtggaatacaccattgaag 652

Db 1243 CTTGGACTAGATTGATTTCCAGGGCGAGGAAATACTCTAGCTTTAAATGGAGAT 1302
Qy 653 catggcagaattgtatttcccttggaaagaaataacccattccaaatttaaatggcgct 712

Db 1303 GGTATCATTTTGTGTTGGATGGGATCAGTCACGTAGCAGTGAACAAATCGCATCTATA 1362
Qy 713 ggtatcatttggatggagacagattgggatacgtcagctcagctcagacaaataata 772

Db 1363 AATTTAGAGTCATGCAAGCTTGGATGGGAAGTGTATACGGAATATGTTAATATG 1422
Qy 773 aattcagaggtaccggaaagggcagtgaggtgggaagtagatataagaaacggcaactatg 832

Db 1423 ATTATTATATGACGCTGATATTGATGATGATCACCAGAGTAGTAATGAATTAAGAA 1482
Qy 833 attaccttatgtatgcagacattgatggatccatccagaagtaatacaatgaacttagaa 892

Db 1483 ATTGGGGTGTGGTACAAACACATTAGATTCGATGATGATTAAGATGATGCGGTTA 1542
Qy 893 attggggagttgtgtatacaaaatacatttaattagattgatttagaatcgatgctga 952

Db 1543 AACATATAAGTATAGCTTTACGCGGATGGGATTAATCATCGTTAGNAGTGCACAGGTA 1602
Qy 953 acataataatacagctatacagagattggctaacacatgctgcgttaacacacagga 1012

Db 1603 AAAATATGTTGGGTTGTGTAGTTTGGAAAGATGATTTAGTGTGCAATTTGAAACATATC 1662
Qy 1013 aaccaatgttgcagatttggaaaataatgaacctgtgcaatgcgaataaactatt 1072

Db 1663 TGCAGAAAACAACTGGAACCATTCAGTCTTTGATGTCGCGTTACATTAATATCTTTATA 1722
Qy 1073 taaataaacaagttggaatcactcctgttcgatgttcctcttcattataattgtaca 1132

Db 1723 ATGCATCAAAAAGCGGGAACATGATGATGGAACATATTTAATGGAAACGTTGTTTC 1782
Qy 1133 atgcataataatggtgggctatttggatagagaataattttcaaatggttctgtgctac 1192
```

---

```
Db 1783 AACGACATCCAAGTCATGCTGTAAACATTTGTTGATAATCATGATTCGCAGCCTGAAGAAG 1842
Qy 1193 aaaaacacctatcatcagtcacattgttgataccatcagctcgcagcgagggaag 1352

Db 1843 CATTAGAAATCTTTTGTGAAGAAATGTTTAAACCATTTAGCGTATGCGCTTACATTAACGC 1902
Qy 1253 catggaaatccttgttcaatcgtgttttcaaacacctgcgcatatgcattgattctgacaa 1312

Db 1903 GTGAACAGGATACCCCTTCTGTATTATTACGGAGATATTATTATGGGATTCACACATCGGAG 1962
Qy 1313 gggagcaaggttacccctccgtattttacggtgattactacggtataccaactcatggtg 1372

Db 1963 TGCAGCAATGAGATCAAAATTCGATTTTGAAGCAGCTGCAAAAGTATGTCATACG 2022
Qy 1373 ttccctcgatgaatactcaaatgtatccactctcgcagcagctcaaacgctatgacctag 1432

Db 2023 GAAACAAATGATTACTTAGACCATCAATATATCATTTGGTGGAGCGCTGAAGGGAATA 2082
Qy 1433 gaaccaacatgatttttgcattcatgatattatcggctggcagcagagagagggaca 1492

Db 2083 CAGCACACCCCAATTCAGGCTAGCTACATCATGCTGTATGAGCGGCTGGAAGTAAAGT 2142
Qy 1493 gctcccaaccacaaattcaggacttgcactattatgctcgatggccagggggtataaat 1552

Db 2143 GGATGTTTGTGGCGTAAATAAGGCTGCTCAAGTATGAGTATGAGTATATACAGGAACCGTA 2202
Qy 1553 ggaatgattcgcggaaacataaaagctggcgaagtatgagagatatccaccggaatacaggt 1612

Db 2203 CAGTACGGTTTACATCAATCAATGACAGCGTTGGGGCAATTTCTCTGTAATGAGGCTCAG 2262
Qy 1613 ctgtaacgctcaccattaatcagatggttgggggaatttcactgtaaacgagggggcag 1672

Db 2263 TTTCTATTGCGTCAACAAATAAAGTGGAA 2293
Qy 1673 ttccggttgggtgaagcaataataagaa 1703

RESULT 2
LOCUS      BSU22045      2447 bp      DNA      BCT      22-MAR-1995
DEFINITION Bacillus sp. alpha-amylose (Amy) gene, complete cds.
ACCESSION  U22045
NID         g722278
KEYWORDS    .
SOURCE      Bacillus sp. (TS-23).
ORGANISM    Bacillus TS-23
            Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
            Bacillaceae; Bacillus.
REFERENCE   1 (bases 1 to 2447)
AUTHORS     Lin, L.-L., Chu, W.S. and Hsu, W.H.
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 2447)
AUTHORS     Lin, L.-L.
JOURNAL     Direct Submission
TITLE       Submitted (01-MAR-1995) Long-Liu Lin, Food Industry Research
INSTRUMENT     Institute, Culture Collection and Research Center, 331 Food Road,
Hsinchu, Taiwan 300, Republic of China
FEATURES    Location/Qualifiers
            source          1..2447
                        /organism="Bacillus TS-23"
                        /db_xref="taxon:38441"
            source          1..2447
                        /organism="Bacillus sp."
                        /strain="TS-23"
                        /db_xref="taxon:1409"
            gene            241..2082
                        /gene="Amy"
            CDS             241..2082
                        /gene="Amy"
                        /EC_number="3.2.2.1"
                        /codon_start=1
                        /transl_table=11
                        /product="alpha-amylose"
```



QY	1007	caggtaaaccaaatgtttgcagttgcagaattttggaaaaatgacctgtgctgcaatcgaaa	1066
Db	1157	ATTACATTACAAAACAAATGGATGCGATGTCATATTATTGATGCACCTTTCATACAACT	1216
QY	1067	actatttaataaacaagaattggaatcactcgcggttcgatgttcctctcatcataatt	1126
Db	1217	TTTATACCGCTTCCAAATCGAGTGGATATTTTTGACATCGCTTATTTTTGAATAATACAT	1276
QY	1127	tgtacaatgatctaataagtggtcattttgatatgagaaaatattttaaatggttctg	1186
Db	1277	TAATGAAGATCAACCTTCACTCGCTGTGACACTTGTGATTAACACGACGACGCAACGAG	1336
QY	1187	tcgtacaaaaaacccctcatcatgtagtcacattgttgataaaccatgactctcagccag	1246
Db	1337	GGCAATCTTTACAGTCATGGTTCGAACCTTGGTTTAAACCACACTTGCATTAGCGCTTTATT	1396
QY	1247	gagaagcatgggaatcccttgttcaatcgtggttcaaacaccactggcatatgcatgtattc	1306
Db	1397	TNACGAGACAAGAGGGATATCTTTCGTGATTTTACGGTGACTATTATGGAATCCCGAAAT	1456
QY	1307	tgacaaggagcaaggttaccccttcgtattttacggtgattactacggtataccaactc	1366
Db	1457	ACAATATTCAGGATTTAAAAGCAAAATCGACCCGCTTTTAAATGTCGTGGGACATATG	1516
QY	1367	atggtgtcttcctgatgaattctaaattgatccactctgcaggcgactcaaacglatg	1426
Db	1517	CCATATGAACACACAGTGATTACATTCACCATCAAGACATATTGATGATGACACGCGAAG	1576
QY	1427	cctacggaacccaacatgattattttgatcatcatgattattcgtgtgacgagagaag	1486
Db	1577	GCAATTGATACAAAACCAAACTCTGGACTGGCGGCTTTAAATTACCGACGGCCCTGGCGGAA	1636
QY	1487	ggacagcctcccaaccacaattcaggacttgcactattatgtcogattgggccaggggta	1546
Db	1637	GCAAATGGATGATGTGCGGTAAAAAACATGCTGAAAAGATATTTTATGATTAACCGGAA	1696
QY	1547	ataaatggtatgtcgggaacataaagctggcccaagtatggagagatatcatcacccgaa	1606
Db	1697	ACCGAAGTCACACAGTAACGATTAATCGGATGCTGGGAGATTTTAAAGTAACCGGAG	1756
QY	1607	ataggctcgtaccctcaccattaatcagatggttgggggaatttctactgtaaacggag	1666
Db	1757	GCTCCGTTTCGATTGGGTG 1776	
QY	1667	ggcaggttccggttgggtg 1686	
RESULT	3		
LOCUS	AF032864	1990 bp	DNA
DEFINITION	Bacillus stearothermophilus alpha amylase (ami)	BCT	25-NOV-1997
ACCESSION	AF032864		
NID	g2642325		
KEYWORDS			
SOURCE	Bacillus stearothermophilus.		
ORGANISM	Bacillus stearothermophilus		
REFERENCE	Eubacteria; Firmicutes; Low G+C gram-positive bacteria;		
AUTHORS	Bacillaceae; Bacillus.		
TITLE	1. (bases 1 to 1990)		
JOURNAL	da Silva,A.C.R., Fernandes,E. and Pueyo,M.T.		
FEATURES	Submitted (03-NOV-1997) Physiology, ICB, Av Prof Lineu Prestes, Sao		
	Paulo, SP, Brasil		
	Location/Qualifiers		
	1..1990		
	/organism="Bacillus stearothermophilus"		
	/strain="ATCC 31195"		
	/db_xref="taxon:1422"		
	/db_xref="ATCC:31195"		
	149..1798		
	/gene="ami"		
gene	149..1798		
CDS	149..1798		
	/gene="ami"		
	/codon_start=1		

```

/transl_table=11
/product="alpha amylase"
/db_xref="PID:g2642326"
/translation="MIFTHRIIRKGMWFLAFLLTASLCTPGQHAKAAAPFNGTMMQ
YFWYLDGGTLTKWTGAENNLISAGITMVLPPAYKGTSDSGVGYDYLDLGEF
NQRGTVRTYKTAQYLQATIAAHAAQMVYADVDFHKGAGDETWDAVEVNPDSR
NORISGTYOIAWTKYDFPCGRGNTYSEFKRWYHFQDVWDSEKLSRIYKRGIGKA
WDEVEVTENGNDYLYADLMDHPVETELKNKGWYVNTNIDGFLDVAVKHIFKS
FFPDWJUSYVRSGTGPDLFTVGEYSVDNINKLHNYITKTNGTSLMSLDAPLHNKFYTSK
SGSAGDMRTLMTLNTMKDQPTLAVTEVDNHDTEPGQALQSWDWPFKPLAYAFLTRQ
EGYCPVGYDYGIGIPOYNTPSLKSIDPLLIARRDYAYGTQHDYLDHSDIIGWTRGV
TEKPGSGSLAALITDGGGSKWMTYVQKHAGKVFYDLTGNRSDTVDLNSDGGWGEFKVG
GSVSVWPKRTIVSTLARPITTPRTPGFEVRWTEPRLVAMP"
529 a 496 c 488 q 477 t

```

Query Match	24.3%	Score 432	DB 12	Length 1990
Best Local Similarity	65.6%	Pred. No. 0.00e+00		
Matches	940	Conservative	0	Mismatches 490; Indels 3; Gaps 3
Db	263	AACGGCACCATGATGACGATTTTGAATGTACTTCCGGGATCATGGCAGCTTATGACAC	322	
QY	253	aatggaaccatgatgacgtatttgaatgagcatcttgcacaaatgacgggaaccactggaac	312	
Db	323	AAAGTGGCCAAATGAAGCCAAACACTTATCCAGCGCTTGGCATCACCGCTCTTTGGCTGCCG	382	
QY	313	aggttacgagatgacgcagcacttaagagataaaggagattaccgcgtttggattcct	372	
Db	383	CCGCGTTACAAAGAACACGCCGACGGAGCTAGGAGTACGGAGTATACGACTTGTAAGAC	442	
QY	373	ccgcgatggaagggaacttcgaaaatgatgttgggtatggtgcctatgatttgcacgat	432	
Db	443	CTCGGGAAATCAATCAAAAAGGACCGTCCGCACAAAAATATGGACAAAAAGCTCAATAT	502	
QY	433	cttgggtgagtttaacccaaaggaaacgcgttcgtacaaaataatggcacaaggagtcagttg	492	
Db	503	CTTCAAGCCATTCAACGGCGCCACGCGCGTGGAAATCCAAAGTCTACGCCGATGCTGTGTTT	562	
QY	493	caagggtgcctgacatctttgaaaaataacgggattcaagtttatgggatgctgtagtc	552	
Db	563	GACCATAAAGGCGCGCTCAGGSCACGGAAATGGGTGAGCGCGTGAAGCTCAATCCGCTC	622	
QY	553	aatcataaagggtggaagcagcgggacagagatggttaaatgcggtgtggaagtgaaccgaagc	612	
Db	623	GACCGCAACCAAGAAATCTCGGCACTATCAAAATCCAAGCATGGACGAAATTTGATTTT	682	
QY	613	aaccgaaccagaataatcagggtgaatacacccattgaagcatggacgaattgatcttc	672	
Db	683	CCGGGCGGGCAACACCTACTCCAGCTTTAAGTGGCGTGTACCATTTTACGCGGCTT	742	
QY	673	cttggaagggaataaccctatcccaactttaatggcgtgtgtacattttgatgggaca	732	
Db	743	GATTGGGACGAAGCCGAAAAA-TTAAGC-CGCAATT-TACAAATTCGGGGCATCGCAAA	799	
QY	733	gattgggatcagtcacgtcagcttcagaacaaaataataaatcagaggtaccggaaag	792	
Db	800	GCCTGGGATGGGAGTAGACACAGAAACGGAAACTATGACTACTTAAATGATGCCGAC	859	
QY	793	gcattgggactgggaagtagatagagaagcgaactatgattagcttatgattgatcgac	852	
Db	860	CTTGATATGGATCATCCCGAAGCTGTGACCGAGCTGAAAAAGCTGGGGAAATGGTATGTC	919	
QY	853	attgatattgatcatccagaagtaataatgaacttagaatactggggagtttggcataca	912	
Db	920	ACACACGAACATTGATGGGTTCCGGGTTGATCCGCTGAACGACTATTAAGTTCAGTTTT	979	
QY	913	aatacacttaactagatgattagaatcgatgctgtgaacaataataatacagcat	972	
Db	980	TTTCTGATTGTTGTGCTGATGTGCTTCTCAGACTGCGACCGCTATTTACCGTCCGG	1039	
QY	973	acgagagattggttaacacatgtgcgttaacccaacaggtaaaccaatgttgcagttgca	1032	
Db	1040	GAATATTGAGCTATGACATCAACAAAGTTGCACAATTACATTCGAAACAAACGGAACG	1099	

QY	1033	gaattttgaaaaatgaccttgcataatcgaaactatttaataaacaagtgtggaat	1092
Db	1100	ATGCTTTTGTGTTGATGCCCGGTTACACACAAATTTTATACCGCTTCCAATACAGGGGGC	1159
QY	1093	cactcogtgcgatgtctctcttcattataattgtacaatgcataataatgctggtgcgc	1152
Db	1160	GCATTGTGATATGCCGACGTTAATGACCATATCTCTCATGAAGATCAACCGACATCGCC	1219
QY	1153	tatttggatgagaaatattttaaattgtctgtcgtaaaaaaacacctatacatgca	1212
Db	1220	GTCACCTCGTTGATAATCATGACACCGAACCGGCCAACGCTGCAGTCATGGGTGCAC	1279
QY	1213	gtcacatttgtataaacatgactctcagcaggagaagcatttggaaatcctttgtctcaa	1272
Db	1280	CCATGGTTCAACCGTTGGCTTACGGCTTTATTCTTAACTCGGCAGGAAGATACCCGTGC	1339
QY	1273	tcgtgggtcaaacacctggcatatgcattgattctgcagaaggagcaggttacccttcc	1332
Db	1340	GTCTTTTATGTGGACTATTATGGCATTTCACAAATATAACATTCTCTTCGTGAAAGCAAA	1399
QY	1333	gtaatttcagggtactactacggtataccaaactctcgtgttctcttcgatgaaatctcaa	1392
Db	1400	ATCGATTCGCTCCTCATCGCGCGCAGGAAATATGCTTACGGAACGCAACATGATTATCTT	1459
QY	1393	attgatccactctcgtcaggcacgtcaaacgtatgcctacggaacccaacatgattattt	1452
Db	1460	GATCACTCCGACATCATCGGTGGACAGGGGAAGGGTCACTGAAAACACGATCCGG	1519
QY	1453	gatacatgatattatcggctcggacgagagaaggggacagctcccaacccaattcagga	1512
Db	1520	CTGGCCGCACGTATCAGCATGGCGCGGAGGAAGCAAAATGGATGTACGTGGCAACAA	1579
QY	1513	cttgcacatttatgtccgatggccagggggtataaaatggatgtatgcggggaacat	1572
Db	1580	CACGCTGAAAAGTGTCTATGACCTTACCGCAACCGGAGTGACACCGCTCACCATCAAC	1639
QY	1573	aaagctgccaaagtatgagagatatcacccggaatagggcttcggaccgtcaccaataat	1632
Db	1640	AGTGATGATGGGGGAATTCAAAGTCAATGCGCGTTCGGTTTCGGTTTGGGT	1692
QY	1633	gcagatggtttggggaaatttcaactgaaacgagggggcaattctcagtttggat	1685

RESULT	4
LOCUS	BACAMYL8 3048 bp DNA
DEFINITION	Bacillus stearothermophilus alpha-amylase gene.
ACCESSION	M11450
NID	g142504
KEYWORDS	
SOURCE	Bacillus stearothermophilus DNA.
ORGANISM	Bacillus stearothermophilus Eubacteria; Firmicutes; Low G+C gram-positive bacteria; Bacillaceae; Bacillus. 1 (bases 1 to 3048) Nakajima,R., Imanaka,T. and Aiba,S. Nucleotide sequence of the Bacillus stearothermophilus alpha-amylase gene J. Bacteriol. 163, 401-406 (1985)
JOURNAL	85234394
MEDLINE	
FEATURES	Location/Qualifiers 1..3048 source

```

CDS
      db_xref="taxon:11422"
      680..12329
      /partial
      /note="alpha-amylase (gtg start codon; EC 3.2.1.1)"
      /codon_start=1
      /transl_table=1
      /db_xref="PID:gl42505"
      /translation="VFTHRIIRKGMFLFLALLFLLFCTQPAKAAAPFGNTWQ
      YFYLDDGTGLTWKVAENALNSLSGLTALWLPAYKTSRSDVGVGVVDLYDLGEC
      NQGVARTYGVKAOYLQIAQAHAAGMOVYADVVDVFDKGGADGETWDAVEVNSDR

```

NOEISGTVQIAWTKFDPGPGNTVSSKPKRWYHEDGVWDSDSKLSRIYKPGIGKA  
WDEVDIENGNDIMLADLDHDEEVTELKSWGKRYVNTINDGFLDUNVAKHIFS  
FPPWMLSDVRSQKLPFVEGYEYSNDKLNKYIMKTNGTSDLEDPALNHYTASK  
SGGTFDMRTLMNLMKQDPTLAVTFVNDHTEPGQALSGWDPKFLAYFLILRQ  
EGPCYVGDYGIQPNILSKLSDHILARRDYAYGQHDYLDHSDILGWTREGV  
TEKPSGLAALITDGPQSGKMYGVQKHAQVYFYDLTGNSRSTVITNSDQGWGFEKVG  
GVSVPKPKTVSTIAWSITVGTDFEFTWTEPVLNVAWP"

BASE COUNT	766 a	795 c	690 g	797 t
ORIGIN	GSVSVWVFRKRIIVSIIAWSTIIIRFWII			

Query Match	24.3%;	Score 432;	DB 12;	Length 3048;
Best Local Similarity	65.6%;	Pred. No. 0.00e+00;		
Matches 940;	Conservative 0;	Mismatches 490;	Indels 3	

Db	794	AACGGCACCATGATGCAGTATTTTGAATGGTACTTCCGGATGATGACGACTTATGGGACC	853
Qy	253	aatgggaccatgatgcagtatgtttgaatggcatttgcataatgcgggaaccactggaac	312
Db	854	AAAGTGGCCAAATGAAGCCACAACACTTATCCAGCCTTGGCATACCCGCTCTTTGGCTGCCG	913
Qy	313	aggttagcagatgcgcagctaaccttaagaagtaaaaggattaccgcgtgttggattccot	372
Db	914	CCCGCTTATAAAGGAACAACGCGCAGGACGTAGGTACGGAGTATACGACTTGTATGAC	973
Qy	373	cctgcaggggaagggacttcgcaaaatgatgttgggcatggtgcctatgatttgcagat	432
Db	974	CTCGGTGAATTCATCAAAAGGGCGCTCGCACAAAATACGGAAACAAAAGTCTCAATAT	1033
Qy	433	cttgtgagttaaccaaaaggaaacgcgtccgtacaaaaataggcaaaagatcagttg	492
Db	1034	CTTCAAGCATTCAAGCGGCCACGCCGCTGTGGAAATGCAAGTGTACGCCGATGTCGTGTTT	1093
Qy	493	caagtgccgtgcacttttgaaaaaataacgggattcaagttattggggatgctcgtgatg	552
Db	1094	GACCATAAAGCGCGCGGACGGACCGAATGGTGGTGACGCCGTCGAAGTCAATCCGTC	1153
Qy	553	aatcataaaggcgaagcagacggacgagatgggtcaaatcggtgggaagcgaaccgaagc	612
Db	1154	GACCGCAACCAAGAAATCTCGGGCACCTATCAAAATCCAAGCATGCGCAAAATTTGATTTT	1213
Qy	613	aaccgaaaccaagaatatcaggtgaatacacaccattgaagcatgagcaaatattgatttc	672
Db	1214	CCCGGGGGGGCACACCTACTCCAGCTTTTAAGTGGCGCTGGGTACCATTTTGTATGGCGTT	1273
Qy	673	cctggaagaggaaatacccatcccaactttaaaatggcgctgggtatcatatttgaeggaca	732
Db	1274	GATTGGGACGAAACCGCAAAA-TTGAGC-CGCATT-TACAAATTCGCGGCATCGGCAAA	1330
Qy	733	gattgggatacgtcacgtcacgttcagaaacaaataataaattcagaggtacocggaag	792
Db	1331	CGGTGGGATTGGGAAGTAGACACGGAAACGGAACCTATGACTACTTAAGTATGCCGAC	1390
Qy	793	gcattggagctgggaagtatatagagaacggcaactatgattaccttatgtatgcagac	852
Db	1391	CTTGATATGGATCATCCGAAGTCGTCGACTGAGCTGAAAAAGCTGGGGAAATGGTATGTC	1450
Qy	853	attgatattgatcatccagaagtaatacaatgaacttagaaaatggggagtttggtataca	912
Db	1451	AACACACGACATGTATGGTTCGGGCTGTAGTCCGCTCAAGCATATATAGTTCAGTTTT	1510
Qy	913	aataccttaactagatgatttagaatgatgtgtaaacatataaataacagctat	972
Db	1511	TTTCTGTATGCTGTCTGATGTCGGTTCTCAGACTGGGAAGCGGCTATTACCGTTGGG	1570
Qy	973	acgagagatggctaaacacatgtgcgttaaacaccacaggttaacccaaagtttgcagttgc	1032
Db	1571	GAATATTGGAGCTATGACATCAACAAGTTGTCACAAATTTACATTTATGAACAAACGGAACG	1630
Qy	1033	gaatttggaaaaatgaccttgcgaatggaaacactatttaataaacaacaagtggat	1092
Db	1631	ATGTCTTTGTTGATGCCCGGTTACACACAAATTTTATACCGGCTTCCAATACAGGGGG	1690
Qy	1093	cactccggttcgattgctccttcatttaaatattgacaaatgcatcatcagtggtggc	1152

[illegible]

RESULT	5	E01181	1680 bp	DNA	PAT	26-NOV-1996
LOCUS						
DEFINITION		DNA encoding highly thermostable alpha-amylase.				
ACCESSION		E01181				
NID		G2169440				
KEYWORDS		JP 1987104580-A/2.				
SOURCE		Bacillus stearothermophilus.				
ORGANISM		Bacillus stearothermophilus				
		Eubacteria; Firmicutes; Low G+C gram-positive bacteria;				
		Bacillaceae; Bacillus.				
REFERENCE		1 (bases 1 to 1680)				
AUTHORS		Kuno,Y. and Akira,S..				
TITLE		DNA-EXPRESSING HEAT-RESISTANT ENZYME				
JOURNAL		Patent: JP 1987104580-A 2 15-MAY-1987;				
		YAMANE KUNIO, HIGETA SHOYU KK				
COMMENT		OS Bacillus stearothermophilus				
		PN JP 1987104580-A/2				
		PD 15-MAY-1987				
		PI 30-OCT-1985 JP 1985241302				
		PF YAMANE KUNIO, SOMA AKIRA				
		PC C12N15/00//C12N9/28;				
		CC strandedness: Single;				
		CC topology: Linear;				
		CC hypothetical: No;				
		CC anti-sense: No;				
		CC *source: strain=A631;				
		CC *source: clone=ptUB616;				
		FH Key Location/Qualifiers				
		FH CDS				
		FT mat_peptide				
		FT Location/Qualifiers				
		FT 1..1680				
		/product='highly thermostable alpha-amyl'				
FEATURES		1..1680				
source		/product='highly thermostable alpha-amyl'				

Db		1102	GTCACTTCGTTGATATCATGACACCGAACCGGCAACCGTGTGACGTGATGGGTGCAC	1161
Qy	1213	gcacatttgatgaataaccatgactcagccaggaagacattggaatcctttgttca	1272	
Db	1162	GCATGGTTCAACCGTTGGCTTACGGCTTTATTCTTAACCTGGGAGGAGATACCCGTGC	1221	
Qy	1273	tcgtggtcaaacacactggcatatgcatgtattctgacaaggagcaaggttaaccttcc	1332	
Db	1222	GTCCTTTATTGAGTATTTATGGCATTCACAAATATAACATTCCTTCGCTGAAAGCAAA	1281	
Qy	1333	gtattttacggtgtattactacggtatacacaactcattggtctcttcctcga tgaatctaaa	1392	
Db	1282	ATCGATCCGCTCCTCATCGCGCAGGGATTATGCTTACGGAACGCAACATGATTATCTT	1341	
Qy	1393	attgattccacttctgagcagctcaagatgcctacggaacccaacatgattatttt	1452	
Db	1342	GATCACTCCGACATCATCGGTTGGGAGGAAGGGTCTACTGAAAACACGAGTCCGGA	1401	
Qy	1453	gatcatcatgattattatcgctggacgagagaaggagacgctcccaacccaatttcagga	1512	
Db	1402	CTGGCGCATTTGATCACCGATGGCGGAGGACGCAAAATGGATGTACGTTCGCAACAA	1461	
Qy	1513	cttgcgaacttatgtctccgatgggcccagggggttaataatcgtatgcatgctgggtacccat	1572	
Db	1462	CACGCGGAAAAGTGTCTATGACCTTACCGCAACCGGAGTGACACCGTCAACATCAAC	1521	
Qy	1573	aaagcttgccaaagtacggagagatatcacgcggaataagctctggtacccgtccaccattaat	1632	
Db	1522	AGTGATGGATGGGGGAATTCAAATCAATGCGCGTTCGGTTTCGGTTTCGGT	1574	
Qy	1633	gcagatggttggggaatttcactgtaaaacggagggcgagtttcggttgggt	1685	
6				
LOCUS	E01180	1719 bp	DNA	26-NOV-1996
DEFINITION	DNA encoding highly thermostable alpha-amylase.			
ACCESSION	E01180			
NID	g2169439			
KEYWORDS	JP 1987104580-A/1.			
SOURCE	Bacillus stearothermophilus.			
ORGANISM	Bacillus stearothermophilus			
REFERENCE	Eubacteria; Firmicutes; Low G+C gram-positive bacteria;			
AUTHORS	Bacillaceae; Bacillus.			
TITLE	1 (bases 1 to 1719)			
JOURNAL	Kunio, Y. and Akira, S.			
COMMENT	DNA-EXPRESSION HEAT-RESISTANT ENZYME			
	Patent: JP 1987104580-A 1 15-MAY-1987;			
	YAMANE KUNIO, HIGETA SHOYU KK			
	OS Bacillus stearothermophilus			
	PN JP 1987104580-A/1			
	PD 15-MAY-1987			
	PF 30-OCT-1985 JP 1985241302.			
	PI YAMANE KUNIO, SOMA AKIRA			
	PC C12N15/00//C12N9/28;			
	CC strandedness: Single;			
	CC topology: Linear;			
	CC hypothetical: No;			
	CC anti-sense: No;			
	CC *source: strain-A631;			
	CC *source: clone-pTUB613;			
	Key			
	Location/Qualifiers			
	1. .1719			
	CDs			
	1. .1719			
	/product-'highly thermostable alpha-amylase'			
	FT mat_peptide			
	1. .1716			
	/product-'highly thermostable alpha-amylase'			
	FT Location/Qualifiers			
	1. .1719			
	/organism='Bacillus stearothermophilus'			
	/db_xref='taxon:1422'			
	BASE COUNT			
	457 a 423 c 434 g 405 t			
	ORIGIN			

Query Match	24.2%;	Score 430;	DB 17;	Length 1719;
Best Local Similarity	65.5%;	Pred. 0.00e+00;		
Matches	939;	Conservative	0;	Mismatches 491; Indels 3; Gaps 5
Db	184	AACGGCACCATGATGCAGTATTTCGAATGGGTACTTGC CGGATGATGCACGTTATGGACC	243	
Qy	253	aaTgggaccatgatgcagtcatttggaaTggcatttgccaaatgcagcggaaccactggaac	312	
Db	244	AAAGTGGCCAATAGAGCCAACCAATTATCCAGCCTTGCATCAACCGTCTTTGGCTCGCG	303	
Qy	313	aggttacgagatgacgcagcctaacttaaagagtaaaaggattaccgcgtgttggaattcct	372	
Db	304	CCGCTTTATAAGGAACAAGCCGACGCGACGTAGGGTACGGAGTATACGACTTGTATGAC	363	
Qy	373	cttgcattgaaaggggacttcgcaaatgatgttgggtatggtgctatgatttgtacgat	432	
Db	364	CTCGGTGAATTCAATCAAAAAGGGCGCTCGCGACAAATACGGAACAAAGCTCAATAT	423	
Qy	433	ctTggTgagtttaaccaaaggaaggaacggtccgtaacaaatatggcaaaaggatcagttg	492	
Db	424	CTTCAAGCCATTCAAGCCGCCACACCGCTGGAATGCAAGTGTACGCCGATTCGTTGTC	483	
Qy	493	caaggtgcgcgtgacatctttgaaaaataaacgggattcaagtttatggggatgctgtgatg	552	
Db	484	GACCATAAAGGGGGCGGACGCGCAGGAATGGGTGGACGCCGTGCAAGTCAATTCGCTC	543	
Qy	553	aatcataaaggTggagcagcagggacagagatTggtaaatgcggtTggaagtgaaccgaagc	612	
Db	544	GACCGCAACCAAGAAATCTCGGCACCTATCAATCCAAAGCATGGACGAAATTTGATTTT	603	
Qy	613	aaccgaacccaagaataTcagtgTgaatcacaccatTgaagcatgagcaagaatttgatttc	672	
Db	604	CCCGGGCGGGCAACACCTACTCCAGCTTTAAGTGGCGCTGTTACCATTTTGTATGGCGTT	663	
Qy	673	ctTggaaaggaaatacccatTccaaactttaaagtgcgtggtatcatctttgatTgggaca	732	
Db	664	GATGGGACGAAAGCCGAAAA-TTGAGC-CGCAAT-TACAAAATCCCGCGCATCGCGAAA	720	
Qy	733	gattTggatcagtcagctcagcttcagaacaaaataataaattcagaggtTaccggaag	792	
Db	721	CGCTGGGATTGGGAAGTACACACGGNAACGGAACTATGACTACTTAATGTATGCCGAC	780	
Qy	793	gcattgggactTgggaagtagatagagaacgcaactatgattaccctatgtatgcagac	852	
Db	781	CTTGATATGATCATCCCAAGTCGTGACTGAGCTGAAAAGCTGGGGAAATGTTATGTC	840	
Qy	853	attgatattgataatccagaagtaataatgaacttagaataattggggagtttgggtataca	912	
Db	841	AACACACGAACATTGATGGTTCGGCTTGATCCGCTCAAGCATATTAAGTTCAGTTTTT	900	
Qy	913	aatacacttaacttagattgattagaatcgatgctgTgaacataattaaatacagctat	972	
Db	901	TTTCCCTGATTGGTTGTCGAATGTGGTCTCAGACTGCGAAGCGCTATTTACGTTGGG	960	
Qy	973	acgagagattTgggttaacatgTgtaacacccaggtTaaacaaatgtttgcagttTgca	1032	
Db	961	GAATATTGGAGCTATGACATCAACAAGTTGCAACATTTACATTTAGAAAACAAACGGAACG	1020	
Qy	1033	gaattTggaaaaaatgaccttgcTgcgaatTcgaaaaactatttaaataaaacaagtTggaat	1092	
Db	1021	ATGTCTTTGTTGATCCCGCTTACACACAAATTTTATACCGTTCCTCAAAATCAGGGGGC	1080	
Qy	1093	cactcgcgtgtcgatgttccctcttattataatttgcataatgcatctaatagTggTggc	1152	
Db	1081	ACATTGTATGCGCAGGTTAATGACCAATACTCTCATGAAGATCAACCAACATTTGGCC	1140	
Qy	1153	tattttgatagaaataattttaaatggttgtgTcgTacaacaaacacccatacatTgca	1212	
Db	1141	GTCACTTCCTTCAATATCATGACACCGAACCCGCCAAGCGCTGCAGTCTATGGTGCAC	1200	
Qy	1213	gtcaacatttgtataacactgactctcgccagggaggaagcatcttggaatcccttTgTcaa	1272	
Db	1201	CCATGGTTCAACCGTTGGCTTACGCCTTTATCTTAACCTCGCGACGAAGGATACCGCTGC	1260	

BASE COUNT	616 a	613 c	559 g	605 t
ORIGIN				
Query Match	24.2%;	Score 430;	DB 12;	Length 2393;
Best Local Similarity	65.5%;	Pred. No. 0.00e+00;		
Matches	939;	Conservative	0;	Mismatches 491; Indels 3; Gaps 3;
Db	727	AACGGCACCATGATGACGATATTTGAATGGTACTTTCGCGGATGATGGCAGCTTATGGACC	786	
Qy	253	aatgggaccatgagcagattttgaaatggcatttgcgaatgacggaaccactggaac	312	
Db	787	AAATGGCCATGAAGCCAAACAATTCACGCTTGGCATCACCGCTCTTTGGTGGCG	846	
Qy	313	aggttacgagatgacgcagtaacttaagatgaagattaccgcgtgttggattcct	372	
Db	847	CCCGCTTACAAGGAACAAAGCGCAGCGACGTAGGTGAGGATATACGACTTGTATGAC	906	
Qy	373	cctgcaggaaggggaacttcgaataatgatgttgggtatgtgcctatgatttgacgat	432	
Db	907	CTCGCGCAATTCATCAAAAAGGGCCCTCGCACAAAATACGGAACAAAAGCTCAATAT	966	
Qy	433	cttggtgagtttaacaaaagggaaccctcgatcaaaaatattggcaagaagagtcagttg	492	
Db	967	CTTCAAGCCATTCAAGCCGCCACCGCCCTGGAATGCAAGTGTACGCCGATGTGCTGTC	1026	
Qy	493	caagtgccgtgacatctttgaaaaataacgggattcaagtttatggggtgtcgtgatg	552	
Db	1027	GACCATAAAGCGCGCGCAGCGACGAATGGTGGGACGCCGTGCAAGTCAATCCGTCC	1086	
Qy	553	aatcataaaggtgagcagcgggacagagatggttaaatcggttggaagtgaaccgaagc	612	
Db	1087	GACCGCAACCAAGAAATCTCGGGCACCTTATCAAAATCCAAAGCATGGACAAAATTTGATTT	1146	
Qy	613	aacgaaaccagaataatcagtggaatacaccattgaagcatggacagaaatttgatttc	672	
Db	1147	CCCGGGCGGCAACACCTACTTCAGCTTTTAAAGTGGCGCTGGTACCATTTTGACGGCGTT	1206	
Qy	673	cctggaagagaaatacaccattccaacttaaatgacgttggtatcatttgatgggaca	732	
Db	1207	GATTGGGACGAAGCCGAAA-TTGAGC-CGCATT-TACAATTCGCGGGCATCGGCAAA	1263	
Qy	733	gatgggaatgctcagctcagctcagacaaaataataataatcagaggtaccggaag	792	
Db	1264	CGGTGGATTGGGAAGTAGACACAGAAACGAACTACTACTTAAATGTATGCGGAC	1323	
Qy	793	gcattggactgggaagtagatatagaaacgcgaactatgattacccttatgtatgcagac	852	
Db	1324	CTTGATATGGATCATCCGGAAGTCGTACCGAGCTGTAACAACTGGGGGAAATGGTATGTC	1383	
Qy	853	attgatattgatcatccagaagtaataatgaacttagaaattggggagtttggtataca	912	
Db	1384	AACACAAGCAATGATGGGTTCCGGCTTCATGCCGTCAAGCATATTAAGTTTCAGTTT	1443	
Qy	913	aatacactaatcagatggattagatcagctgtgaaacatatataacatcagctat	972	
Db	1444	TTTCCTGATTGGTGTGCTATGTGCGTTCTCAGACTTGGCAAGCGCTATTTACCGTCGGG	1503	
Qy	973	acgagagattggcctaacaacagtgcgttaaacaccacaggttaaaccaattgttcagttgca	1032	
Db	1504	GAATATTGGAGCTATGACATCAACAAGTTGCACAAATTACATTTACGAAACAAACGGACG	1563	
Qy	1033	gaatttggaaaaatgacctgtcgtcaatcgaaacactatttaataaaacaagttggaat	1092	
Db	1564	ATGTCTTTGTTGATGCCCGTTACACAACAATTTATCTGCTTCCAAATCAGGGGGC	1623	
Qy	1093	cactccggtgtcagatgttcctctcattataatttgtaaatgcatctaatagtggtgac	1152	
Db	1624	GCATTTGATATGCGCAGGTTAATGACCAATACTCTCATGAAGATCAACCCGACATTTGGCC	1683	
Qy	1153	tattttgatagaaaatttttaattggttctgtcgtacaaaaaaccccttatcatgca	1212	
Db	1684	GTCACCTTCGTGTATATCATGACACCGAACCCCGGCCAACCGCTGTCATGGTCCGAC	1743	

D	b	510	CTTCAAGCATTCAAGCCGCCACCGCGTGGATGCAAGTGTACGCCGATGCTGCTTC	569
Q	Y			
D	b	493	caaggtgcgtgacatcttttgaaaaataacggatctcaagtttatgggattgctgtagt	552
Q	Y			
D	b	570	GACCAATAAGGGGGGCTCAGCGCAGGAATGGGTGGACCGCTCGAAGTCAATCGCTCC	629
Q	Y			
D	b	553	aatcaaaaggtggagcagacgggcagagatggtaaatcggtgggaagtgaaacccaagc	612
Q	Y			
D	b	630	GACCGCAACCAAGAAATCTCGGGCACCTTATCAAAATCCAAGCATGAGCAAAATTTGATTTT	689
Q	Y			
D	b	613	aaccgaaccaagaatacaggtgaatacacaccattgaagcatggacgaataattgatttc	672
Q	Y			
D	b	690	CCGGGGGGGGCAACACTACTTCAGCTTTTAAGTGGCGCTGGTACATTTTTCAGCGGCTT	749
Q	Y			
D	b	673	cttggaaagggaataaccocatcccaactttaaattggcgtggtatcatttttgatggaca	732
Q	Y			
D	b	750	GATTGGGACGAAGCGGAAAA--TTGAGC--CGCATTT--TACAAATTCGCGGCATCGGCAAA	806
Q	Y			
D	b	733	gattgggatacgttcacgttcagcttcagacaacaaataataatccagaggcaccggaag	792
Q	Y			
D	b	807	CGCTGGGATTGGGAAGTAGACACGGAAACGAAACTATGACTACTTAAATGTATGCGCAC	866
Q	Y			
D	b	793	gcatggactgggaagtagatagagaacggcaactatgattaccattatgtatgcagac	852
Q	Y			
D	b	867	CTTGATATGATCAATCCGAAGTCTGTGACCGAGCTGAAAACTGGGGAAATGGTATGTC	926
Q	Y			
D	b	853	attgatgtgatacatocagaataatcaatgaacttagaaattgggagtttgggtataca	912
Q	Y			
D	b	927	AACACAACGAACATGATCGGTTCCGGCTTGATCGCTCAAGCATATTAAGTTTCAGTTT	986
Q	Y			
D	b	913	aatcacacttaactagatggattagaatcgatgctggaaacataataatacagctat	972
Q	Y			
D	b	987	TTTCTCGATTGGTTGCTGATGGGTTCTCAGACTGCCAAGCGCTATTTCACGTCGGG	1046
Q	Y			
D	b	973	acgagagattgctaacacatgtgcgtaacaccacaggtaaaccaattgttcagttgca	1032
Q	Y			
D	b	1047	GAAATTTGGAGCTATGACATCAACAAAGTTGCACAATTCATTCAGAAAAACAGAGCAAGC	1106
Q	Y			
D	b	1033	gaattttgaaaaatgaccttgcgtgcaatcgaaaaactattcaataaaaaacaagtggaa	1092
Q	Y			
D	b	1107	ATGCTTTGTTTATGATCGCCGGTTACACAACAATTTTATACCGCTTCCAAATCAGGGGGC	1166
Q	Y			
D	b	1093	caatccgtgttgatgttccctctcatataatttgtacaatgcatataatagtggtgac	1152
Q	Y			
D	b	1167	GCATTTGATATGGCAGCTTAATGACCAATACTCTCAAGATCAACCGACATTTGGCC	1226
Q	Y			
D	b	1153	tatttggatatgagaaatatttaaattggtctgcgtacaaaaaacacctatacatgca	1212
Q	Y			
D	b	1227	GTCACTTCGTTGATTAATCATGACACCGAACCCGCCAAGCGCTCGACTCATGGGTGCAC	1286
Q	Y			
D	b	1213	gtcacatttgtataaccatgactctcagccagagagcatggaaatccttctgttcaa	1272
Q	Y			
D	b	1287	CCATGGTTCAACCGTTGGCTTACGCCCTTATTCTTAACCTGGCAGGAGGATACCCGTGC	1346
Q	Y			
D	b	1273	tcgtggttcaaacacctgcatatgatttctgacagggagcaggtttacccctcc	1332
Q	Y			
D	b	1347	GTCTTTTATGGTCACTATTATGGCATTCACAATAATAACATTTCTGCTGAAAAGCAAA	1406
Q	Y			
D	b	1333	gtattttcacggtgattactacaggtataccaaactatggttctcttggatgaactctaaa	1392
Q	Y			
D	b	1407	ATCGATCCCGTCTCTATCGCGGGCAGGGATTATGCTTACGGAACGCAACATGATTATCTT	1466
Q	Y			
D	b	1393	atlgatccactctgcagcagcgtcaaacgtatgcctacggaacccaacatgattatttt	1452
Q	Y			
D	b	1467	GATCACTCCGACATATCGGGTGGACAGGGGAAGGGCGACTGAAAAACAGAGATCCGGA	1526
Q	Y			
D	b	1453	gatcatcatgatattcggcttggacgagagaggggacagctccaccaccaatctcagga	1512
Q	Y			
D	b	1527	CTGGCGCACTGATACCGATGGCGGGGAGGAACCAATGGATGCTAGCTTGGCAACAA	1586
Q	Y			
D				

Oy 1573 aaagctggccaaagtatgagagatatcacgcgaataaggctgttacgcgtaccattaat 1632  
| ||||| |||| | || | ||||| || || || ||||| ||||| ||  
Db 1647 AGTGAATGATGGGGGAATTCAAAGTCAATCGCGTTTCGGTTTCGGTTTGGGT 1699  
||||| ||||| |||| | || | ||||| || || || ||||| ||||| ||  
Oy 1633 gcagatggttgggggaatttcactgtaaacgagcgagttcgcgttggttgggt 1685  
| ||||| ||||| |||| | || | ||||| || || || ||||| ||||| ||

RESULT 9 BACAMYABS 1891 bp DNA BCT 24-OCT-1990  
LOCUS B.stearothermophilus alpha amylase gene, complete cds.  
DEFINITION M57457  
ACCESSION g142481  
NID  
KEYWORDS alpha-amylase.  
SOURCE Bacillus stearothermophilus (library: ATCC 12980) DNA.  
ORGANISM Bacillus stearothermophilus  
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;  
Bacillaceae; Bacillus.

REFERENCE 1 (bases 1 to 1891)  
AUTHORS Suominen,I., Karp,M., Lautamo,J., Knowles,J.K.C. and Mantsaela,P.P.  
TITLE Thermostable alpha amylase of Bacillus stearothermophilus: Cloning,  
expression, and secretion by Escherichia coli  
(in) Chaloupanka J. and Krumpalanzl V. (Eds.);  
JOURNAL EXTRACELLULAR ENZYMES OF MICROORGANISMS: 129-137;  
Plenum Press, New York (1987)

FEATURES Location/Qualifiers  
source 1..1891  
/organism="Bacillus stearothermophilus"  
/db\_xref="taxon:1422"  
/tissue\_lib="ATCC 12980"  
-35\_signal 89..94  
/gene="alpha-amylase"  
/note="P1"  
gene 89..1881  
/gene="alpha-amylase"  
-10\_signal 114..119  
/gene="alpha-amylase"  
/note="P1"  
-35\_signal 185..190  
/gene="alpha-amylase"  
/note="P2"  
-10\_signal 207..212  
/gene="alpha-amylase"  
/note="P2"  
sig\_peptide 235..336  
/gene="alpha-amylase"  
CDS 235..1884  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase"  
/db\_xref="PID:g142482"  
/translation="MLTFHRIIRKGNWFLLAFLTALLFCPTGPAKAAAPFNCTMMQ  
YFENLPDDGTGLTKAQAANNLSSLGALITWLPPAYKYGRSDVGVVDLYDLGEFE  
NQKGAVTGTAQYQLQAIQAHAAGQMAYADVFDHKGGADGETWDAVEVPNSPR  
NQETSGIQIQAATKFDPGRGNTYSFKRWDFDGDWDSESKLSRIYKFRIGIKFS  
WDWEVDENGNDYLMYADLDMDHPVVTELKAGAKWYVNTRTLDGRLDAVKHIKS  
FFPDMLSVSQSCKPLFTVCYESYDIKLHNXYIMKTNGTMSLFDAPLHNKEVTASK  
SGGTEDMTLTMTLMDKQTLAVTFVDNHDTEPGALQSWDFWPKFLAYAFILTRQQ  
EGYPCVFYGDYGIPQNIPLSKIDPLIARDYAQTQHDLDHSDLIIGTWREGV  
TERKPSGALAALITDPPGGSKWMMYVGKHAGKVFDLTGNRSDTVTINSDGWGEEKFWK  
GSVSVWPRTKTVTSTIANSTTPWTDEFWRVTEPRLVAMP"  
337..1881  
mat\_peptide /gene="alpha-amylase"

BASE COUNT 498 a 469 c 464 g 460 t  
ORIGIN

Query Match 24.1%; Score 428; DB 12; Length 1891;  
Best Local Similarity 65.5%; Pred. No. 0.00e+00;  
Matches 938; Conservative 0; Mismatches 492; Indels 3; Gaps 3;

Db 349 AACGCCACCATGATGCAGTATTTGAATGGTACTTCGCCGGATGATGGCACGTATGGACC 408  
| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
Ov 253 aatgacaccatgatgcactattttgaatggcatttgcacaatgccgaacacctgggac 312  
| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||



Db 409 AAGTGCCCAATGAAGCCACAACTATTACAGCCTTGCCATFACCGCTCTTTGGTCGCC 468  
Qy 313 aggttacgagatgacgagcttaactaaagagtaaaaggtattaccgtgttggtatccct 372  
Db 469 CCGCTTATAAAGGAACAACCGCAGCGAGCTAGGTACGGATACGAGTATACGACTTTGTATGAC 528  
Qy 373 cctgcatggaaggggacttcgcaaaatgagtgttggtgctgctgctgattgattgaagat 432  
Db 529 CTCGGTGAATTCAATCAAAAAGGGCGCTCCGCACAAAATACGGAACAAAGCTCAATAT 588  
Qy 433 ctgggtgagtttaaccaaaggaacogtccgtacaaaatattggcaacaagagtcagttg 492  
Db 589 CTTCAAGCCATTCAAGCGCGCCAGCGCTGGGAATGCAAGTGTACGCCGATGTCGTTC 648  
Qy 493 caaggtgctgacatctttgaaaaataacgggattcaagtttatgggtggtgctgctgatg 552  
Db 649 GACCATAAAGCGCGCGCAGCGCACGGAATGGTGTGACGCCGCTCGAAGTCAATCCGTC 708  
Qy 553 aatcataaaggtggagcagacgggacagagatggttaaatcggtggaagtgaaccgaagc 612  
Db 709 GACCGCAACCAAGAAATCTCGGCATCTATCAATCAAGCATGCGAGAAATTTGATTTT 768  
Qy 613 aaccgaaccaaagaatatcaggtgaatacacccattgaagcatgagcaaaatttgatttc 672  
Db 769 CCGCGCGGGGCAACACCTACTCCAGCTTTAAGTGGCGCTGGTACCAATTTTGATGGCGTT 828  
Qy 673 cctggaagaggaataaccattccaactttaaatggcgtggtatcattttgatgggaca 732  
Db 829 GATTGGGAGCAAGCGGAAAA-TTGAGC- GGCATT-TACAAATTCGCGCATCGCCAAA 885  
Qy 733 gattgggacagtcacgtcagcttcagaaacaaatatataatactcagaggtacgggaaag 792  
Db 886 GCGTGGATGGGAATAGACACGGAACGAAACGAACTATGACTACTTAATGTATGCGGAC 945  
Qy 793 gcatggtactgggaagtagatagagaacggcaactatgattacatttatgtatgcagac 852  
Db 946 CTGTGATGATGATCCCGAAGTCGTGACTGACGTAAAGCTGGGCTAAATGGTATGTC 1005  
Qy 853 atgcatgagatggtcgaacacatgtgcgttaacacacaggttaaaccaattggtgtataca 912  
Db 1006 AACACAAGAACACTTGATGGTTCGCGCTTGATCGCTCAAGCATATTAAGTTACGTTTT 1065  
Qy 913 ataacaactaatcagatgattagaatcagatcgtgtaaacattatataacacagctat 972  
Db 1066 TTTCCTGATGTTGCTGATGTCGTCTCAGACTGGCAAGCGCTATTTACCGTTGGG 1125  
Qy 973 acgagagattggcgaacacatgtgcgttaacacacaggttaaaccaattggtcagttgca 1032  
Db 1126 GAATATTGGAGCTATGACATCAACAGTTGCACAAATTACATTATGAACAAACGGAACG 1185  
Qy 1033 gaatttggaaaaatgacctgtgctgcaatcgaaaaactatttaataaaaaacaagttggaat 1092  
Db 1186 ATGTCTTTTGTGATGCCCGCTTACACACAAATTTATACCGCTTCCAAATCAGGGGCG 1245  
Qy 1093 cactcgtgttcgagttcctcttcattataattgtacaaatgcatactaatagtggtggc 1152  
Db 1246 ACATTGATATCGGACGTTAATGACCAATCTCTCATGAAAGATCAACCAACATTGGCC 1305  
Qy 1153 tatttgatagagaataatttaaatgggtctgtctgtaaaaaaacacccctataatgca 1212  
Db 1306 GTCACCTTCGTTGATAATCATGACACCGAACCAGCCGCAAGCGCTGCAGTCATGGGTGCGAC 1365  
Qy 1213 gtcaaatgttga taacctgactctcagccagagagaagcattggaatccctttgtcaa 1272  
Db 1366 CCATGTTTCAACCGCTTACGCTTTTACGCTTTTATCTAACTTCGGCAGGAAGTACCGGTGC 1425  
Qy 1273 tcgtggttcaaacacctggcatgattgattcgtgacaagggagcaaggttacccttcc 1332  
Db 1426 GTCCTTTTATGGTACTATTATGGCAATTCACAAATATACATTCTTCGTGTGAAAGCAAA 1485  
Qy 1333 gtaatttcaggtgattactaggtgataccaactcactggtgttccttcgtgagaatctaaa 1392

Db 1486 ATCGATCCGCTCCTCATCCGCGCAGGAGTATGCTTACGGAACCAACATGATTATCTT 1545  
Qy 1393 attatccacttctgcagcgacgttaaacgctatgctctacggaacccacatgattatttt 1452  
Db 1546 GATCAGCTCCGACATCATCGGTGTGACAAAGGAGGGTCACTGAAAAACCCAGGATCCGGA 1605  
Qy 1453 gatcatcatgattatttcggtggtgacgagagaggggacagctccacccaatttcagg 1512  
Db 1606 CTGCGCCCATGTGATCACCAGTGGCGCGGAGGAGGAACCAATGATGTAGCTTGGCAACAA 1665  
Qy 1513 cttgcaactattatgctcgatggcgagggttaataaatgattgattgctcggaacat 1572  
Db 1666 CACCCCGGAAAAGTGTCTATGACTTACCGGCAACCGAGTGACACGTCACCATCAAC 1725  
Qy 1573 aaagctggccaagtatggagagatatcccggaataagtggtgacctaccattaat 1632  
Db 1726 AGTATGATGGGGGAATTCAAAGTCAATGCGGCTTCGGTTTCGGTTTGGGT 1778  
Qy 1633 gcagatggttgggggaatttcactgtaaacggaggggcgagttcgttgggttgggt 1685

## RESULT 10

LOCUS BSAMYL1 2169 bp DNA BCT 21-MAR-1995  
DEFINITION Bacillus stearothermophilus gene for alpha-amylase.  
ACCESSION X02769  
NID g39799  
KEYWORDS amylase; amylase-alpha; inverted repeat; signal peptide.  
SOURCE Bacillus stearothermophilus.  
ORGANISM Bacillus stearothermophilus  
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;  
Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 2169)  
AUTHORS Ihara,H., Sasaki,T., Tsuboi,A., Yamagata,H., Tsukagoshi,N. and Uda,K.S.  
TITLE Complete nucleotide sequence of a thermophilic alpha-amylase gene: homology between prokaryotic and eukaryotic alpha-amylases at the active sites  
JOURNAL J. Biochem. 98 (1), 95-103 (1985)  
MEDLINE 86008166  
REFERENCE 2 (bases 1 to 2169)  
AUTHORS Tsukagoshi,N.  
TITLE Direct Submission  
JOURNAL Submitted (03-SEP-1985) to the EMBL/GenBank/DBJ databases  
COMMENT Data kindly reviewed (03-SEP-1985) by Tsukagoshi N.  
FEATURES  
Location/Qualifiers  
1..2169  
/organism="Bacillus stearothermophilus"  
/db\_xref="taxon:1422"  
66..74  
/note="inverted repeat a"  
79..87  
/note="inverted repeat a"  
142..150  
/note="inverted repeat b"  
161..169  
/note="inverted repeat b"  
209..214  
/note="put. -35 region"  
233..238  
/note="put. -10 region"  
253..256  
/note="pot. Shine-Dalgarno region"  
306..310  
/note="pot. Shine-Dalgarno region"  
338..343  
/note="pot. Shine-Dalgarno region"  
345..348  
/note="pot. Shine-Dalgarno region"  
354..2003  
/note="put. precursor"  
/codon\_start=1  
/transl\_table=11  
/db\_xref="FID:g580825"



Query Match	24.1%	Score 428;	DB 12;	Length 2169;
Best Local Similarity	65.5%;	Pred. No. 0.00e+00;		
Matches	938;	Conservative	0;	Mismatches 492; Indels 3; Gaps 3;
Db	468	AACGGCACCATGATGCAGTATTTTGAATGTGTACTTTCGCCGATGATGCCAGCTTATGAC	527	
Qy	253	aatgggaccatgagtcgcggtatttgaatggcatttgcgaatgacgggaaccactggaac	312	
Db	528	AAATGGCCCAATGAAGCCAAACAACTTATCCAGCCTTGGCATCACCGCTCTTTGGTGC	587	
Qy	313	aggttcagcagatgcgcagctaaacttaagagtaaaaggattaccgcgtgttggattcc	372	
Db	588	CCCGCTTACAAGGAACAAGCCGACGAGGTAGGGTACGGAGTATAGCACTTGTATGAC	647	
Qy	373	cctgcattggaagggaacttcgcaaatgatgttggtgatgtgctatgatattgtacgat	432	
Db	648	CTCGGCGAATTCATCAAAAGGGACGCTCGGCACAAAATATGGAACAAAAGCTCAATAT	707	
Qy	433	cttggtagtttaaccaaaagggaacgcctcgctacaaaatggaacaaaggatcagttg	492	
Db	708	CTTCAAGCGATTCAAGCGCGCCACGCGCTGGAAATGCAAGTGTACGCGCATGTCTGTTT	767	
Qy	493	caagtgccgfgacatcttggaaaaataaaacgggattcaagtttatgggatgctgcatg	552	
Db	768	GACCATAAAGCGGGCTGACGGCACGGAATGGGTGTACGCCCTGTGAAGTCAATCCGTCC	827	
Qy	553	aatcataaagtgagcagacgggacagagatggtaaatcggtgggaagtgaaccgaagc	612	
Db	828	GACCGCAACCAAGAAATCTCGGGCACCTATCAAAATCAAGCATGACGCAAAATTTGATTTT	887	
Qy	613	aaccgaaaccagaataatcaggtgaaatacaccattgaagcagtcggaatttgatttc	672	
Db	888	CCCGGGCGGGCAACACCTACTCCAGCTTTTAAGTGGCGCTGGTACCACTTTTGA	947	
Qy	673	cctggaagagaataaccattccaactttaaatggcgtggatcatatttgatgggaca	732	
Db	948	GATTGGGACGAAGCCGNAAAA-TTAAAGC-CCGATT-TACAAATTCGCGGCATCGGCAAA	1004	
Qy	733	gattgggatacagtcacgtcagcttcagaacaaaataataaaattcagaggtaccgcgaag	792	
Db	1005	CGGTGGGATTTGGGAAGTAGACACAGAAACGGAACATATGACTACTTAAATGTATGCGGAC	1064	
Qy	793	gcaTgggactgggaagttagatagagaacggcaactatgattacccttatgtatgcagac	852	
Db	1065	CTTGATATGGATCATCCGGAAGTCGTGACCGAGCTGAAAAACCTGGGGAATGGTATGTC	1124	
Qy	853	attgatatgcatccagaagaataatcaatgaacttagaaattgggagtttggtataca	912	
Db	1125	AACCAACGAACATTGATGGGTTCGGCTTGATGCCGTCAAGCAATATTAAGTTCAGTTTT	1184	
Qy	913	aatacacctaatcttagatggatttagaatcatgctgtgaaacataataaacacgatat	972	

PC	C12R1:19),(C12N9/52,C12R1:19);
CC	strandedness: Double;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
FH	Key
FH	Location/Qualifiers
FT	CDS
FT	1. .>1647
FT	Location/Qualifiers
FT	1. .1650
FT	Location/Qualifiers
FT	/organism="Bacillus stearothermophilus"
FT	/db_xref="taxon:1422"
FT	439 a 413 c 419 g 379 t
Query Match	23.9%; Score 424; DB 17; Length 1650;
Best Local Similarity	65.3%; Pred. No. 0.00e+00;
Matches	936; Conservative 0; Mismatches 494; Indels 3; Gaps 3;
Db	115 AACGCGACATGATGCAGCTATTTTGAATGGTACTTTCGCCGATGATGGCGACGTTATGGACC 174
Qy	253 aatgggaccatgatgcagctattttgaatggcaatttggcaatggcgaatgcgggaaccactcggaac 312
Db	175 AAATGGCCAATGAAGCCAAACAACTTATCCAGCCTTGGCATCACCGCTCTTTGGCTGGCG 234
Qy	313 aggttacgagatgcgcagctaaacttaagagtaaaaggattaccgcgtttggattcct 372
Db	235 CCCGCTTATAAAGNACAAAGCCGACGAGCTAGGGTAGCGAGTATACGACTTGTATGAC 294
Qy	373 cctgcataaggagggaacttcgaatgatgttgggtatggtgcctatgattgtacgat 432
Db	295 CTCGGCGAATTCATCAAAAGGAGGACCGCTCGCACAAAATACGGAAACAAAGTCAATAT 354
Qy	433 cttggtgagtttaaccaaagggaacccgcgtacaaatatggcacaagagtcagctg 492
Db	355 CTTCAAGCCATTCAAGCGCCACGCCCTCGAATGCAAGTGTACGCGGATGTCGTGTTC 414
Qy	493 caaggtgcgctgacatctttgaataaatacgggattcaagtttatgggagtgcgtgatg 552
Db	415 GACCATAAAGCGCGCGACGCGAGAAATGGGTGGAGCGCGTGCAGAGTCAATCCGTC 474
Qy	553 aatcataaagtgtagcagcggacagagatggttaaatgcggtggaaagtgaaccggaagc 612
Db	475 GACCGCAACCAAGAAATCTCGGGACCTTATCAATCCAAAGCATGACGACAAATTTGATTT 534
Qy	613 aaccgaaccaagaataatcaggtagaacaccattgaagcatggacgaatttgatttc 672
Db	535 AACGGGGGGCAACACTACTCCAGCTTTAAAGTGGCGTGGTACCATTTTATGACGGCGTT 594
Qy	673 cctggagagaaataccattccacttaaatggcgtgattcatttggatgggaca 732
Db	595 GACTGGGACGAAACCGGAAAA-TTAAGC-CGCATT-TACAAATTCGGCGGATCGGCAAA 651
Qy	733 gattgggatcagtcacgcagcttcagaacaaaataataaattcaggaggtaccgggaaa 792
Db	652 GCGTGGGATTGGGAAGTAGACACGAGAAACGAAACATGACTACTTAATCTATGCCGAC 711
Qy	793 gcatgggactgggaagtagatagagaacggcaactgattacccttatgtatgcagac 852
Db	712 CTTGATATGGATATCCGGAAGTCGTGACCGAGCTGAAAAAACTGGGGAAATGGTATGTC 771
Qy	853 attgatattgatcattcagaagtaatacaatgaactagaattgggaggttggataca 912
Db	772 AACACAGCACATTTGATGGTTCGGCTTGATGCTGTCAAGCATATTAAGTTTCAAGTTT 831
Qy	913 aatacatttaactagatggatttagaatcgatcgtgtgaacataataataacagcatat 972
Db	832 TTTCTGATTTGGTTGTCGTATGTCGGTTCTCAGACTGCGAAGCGCGCTATTTTACCGTCGG 891
Qy	973 acgagagattggctaacacatctgctgtaaacaccacaggttaaaccaatgcttgcagctgca 1032
Db	892 GAATATTTGGAGCTATGACATCAACAAGTTGCACAAATTACATTTACGAAACAAACGGACG 951

BASE COUNT  
 ORIGIN  
 540 a 506 c 512 g 508 t  
 36 bp upstream of Ahai11 site.  
 TGGSVVVRKRTTITSTWPPITVTPGFEFVWTEPRLLVMP"  
 KSPVSGGLAAITDGGSGKMWYTKQHGAKVFYDITGNRSDDTITNSDQGWGEFFKNG  
 EGPCVFYIGYIGPIQNPISLAKSDIDPLLIARDYAYTGHDLDWSPDLIAGWTRGV  
 SGGAFMSYMLNNMLNKRPDIAVTFNDHDETPGQALQASWDVDFKSPALIAILTRQ  
 FPDWILSYVSQSGKELFVYGEVWSYDINKLHNYITKTNGTMSLFDPAPLHNKZYATSK  
 WDFEVTENGNYDLAMVADMDPEWVTEFLKNGKVVYNTNIDGFRDVAADYKIFES  
 NQFISTQYIQAWTKFDNCRGNGTYTSEFKWRYWFDQGVDESKRLSRIRYKPFGRICK

Query Match 23.9%; Score 424; DB 12; Length 2066;  
Best Local Similarity 65.3%; Pred. No. 0.00e+00;  
Matches 936; Conservative 0; Mismatches 494; Indels 3; Gaps 3;

Db	310	AACGGCACCATGATGCAGTATTTTGAATGGGTACTTTCGCCGGATGATGCCACGCTTATGGAAC	369
QY	253	aatgggaccatgatgcagtatttgaatggcaatttggccaaatgacgggaacccatggaac	312
Db	370	AAAGTGGCCAATGAAGCCAACAACTTATCCAGCCTTGGCATCACCGCTTCTTTGGCTGCGG	429
QY	313	aggttacgagatgacgcagctaacttaagagtaaaaggattaccgcgtgtttggatccct	372
Db	430	CCCGCTTATTAAGGAACAACGCGACGACGTAGGGTAGCGAGTAGACACTTATGTATGAC	489
QY	373	ctctgcatgaaagggaacttcgcaaaatgatgttggtatgtgctatgatttgtacgat	432
Db	490	CTCGGCGAATTCNAATCAAAAGAGGACGCTGCCCAAAATACGGACAACAAAGCTCAATAT	549
QY	433	ctctgttgagttaaccaaaagggaacccgtccgtacaaaatatggcacaaggagtcagttg	492
Db	550	CTTCAAGCCATTCAAGCCGCCACGCCGCTGGAATGCAAGTGTACGCCGAGTGTGCTGTC	609
QY	493	caagggtgcogtgacatctttgaaaaataacgggattcaagtttatggggatgtcgtgatg	552
Db	610	GACCATAAAGGCGCGCCGACGCGAAGTGGGTGGACGCGTGGAAAGTCAAACTCCGTCC	669
QY	553	aatcataaagggtggagcagacgaggcagagatggttaatgcggtgggaagtgaacccgaagc	612
Db	670	GACCCCAACCAAGAAATCTCGGCACTTATCAAAATCCAAAGCATGGACGAAATTTGATTTT	729
QY	613	aaccgaaaccaaataatcaggtgaatacacaccattgaagcatggcagaaatttgatttc	672
Db	730	AACGGGGGGGCAACACCTTACTTCCAGCTTTAAGTGGCGCTGGTACCATTTTTACGGCGGT	789
QY	673	cttggaaggaggaatacccattcccaactttaaattgaatgcgtgtatcatcttggatgggaca	732
Db	790	GACTGGGACGAAACCGCAAAA-TTAAAGC-CGCAAT-TACAATTCGGGGCATCGGCAAA	846
QY	733	gattggggtcagtcacgctcagcttcagaaacaaaataataattcagaggccacggaaaag	792
Db	847	GCCTGGGATTGGGAAGTAGACACGGAACGAAACACTATGACTACTTAATGTATGCCGAC	906
QY	793	gcattggactgggaagtagatatagaaacgcaactatgattaccttatgtatgtacagac	852
Db	907	CTTGATATGATCATCCGGAAGTGTGACCGGAGCTGAAAACTGGGGGAAATGGTATGTC	966
QY	853	atigtatggtatcatccgaagaataccaatgaacttagaataatggggagtttgggtataca	912
Db	967	AAACACAACGACATTGATGGGTTCGGCTTGATGCTGTCAACCATATTTAAGTTCAGTTTT	1026
QY	913	aataccttaactatgattgaattagaatcgtatgctgtgaaacatatataacacagtat	972
Db	1027	TTTCCCTGATTGTGTGTATGTGGCTTCTCAGACTGGCAAGCCGCTATTTACCGTGGG	1086
QY	973	acgagagattggttaacacatgttggttaaacaccacaggttaaccaaattgtgcagttgca	1032
Db	1087	GAATATTGGAGCTTACACATCAACAGTTTGCACAAATTACATTACGAAACAAACGGGAACG	1146
QY	1033	gaattttgaaaaaatgaccttgctgcatacgaataactatttaataataaacaaagtggaaat	1092
Db	1147	ATGCTCTTTGTTGATCGCCCGTTTACACAACAATTTTATCTGCTTCCAAATCGGGGGGGC	1206
QY	1093	cactccggtgttcgatttccctctcatataattttgacaaatgcgactaaatgattgtgtgac	1152

```

RESULT 13
LOCUS A47677 1539 bp DNA 07-MAR-1997
DEFINITION Sequence 1 from Patent WO9535382.
ACCESSION G47677
NID g2301618
KEYWORDS
SOURCE Bacillus licheniformis.
ORGANISM Bacillus licheniformis
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 1539)
AUTHORS Van, D.L. and Ahhle, W.
TITLE NOVEL AMYLOLYTIC ENZYMES DERIVED FROM THE B. LICHENIFORMIS alpha
-AMYLASE, HAVING IMPROVED CHARACTERISTICS
JOURNAL Patent: WO 9535382-A 1 28-DEC-1995;
GIST BROCADES BV (NL)
COMMENT Other publication AU 2524795 960115.
FEATURES
source 1..1539
/organism="Bacillus licheniformis"
/strain="CBS407.83"
/db_xref="taxon:1402"
1..87
sig_peptide 1..1539
CDS

```

sig\_peptide  
CDS

```

/ascii_table=11
/transcript_table=11
/db_xref="PID:e306281"
/db_xref="PID:g2301619"
/translation="MKQKQYVIRARLLTLLFALFLPHSAAAAANGLTQMIFYEWTM
PNDQHWKRLQNSAYLAEGHITAVIIPAYKQTSQADGYGAYDLDIGEFHQKTV
RTYKTKGASAKLSHRSINVTYGDVIVNHKGADATEDYATVDEPADNRNVTSG
EHLIKATWHEFGPGRSISYDFKWHYHFDGFDGDESKRLNRYKFGQKADWEVSE
NNDYLYLADIDVDHPDVAEIKRWGTYVANELODGFPLDPAVKHIKTFSLRDWNV
VRETKGEMETVAEDVDQAGALENTNKTNEVSDVDFVPHFAASTOGGSDYWRH

```



[illegible]

Job time : 2996 secs.

[illegible]